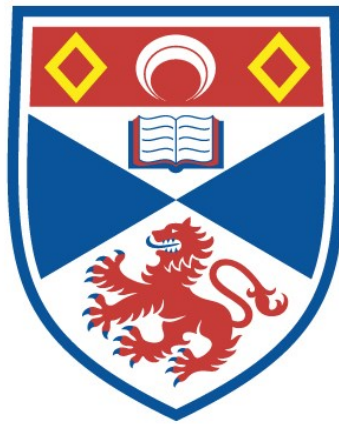


BAYESIAN ANALYSIS FOR QUANTIFICATION OF  
INDIVIDUAL RAT AND HUMAN BEHAVIOURAL PATTERNS  
DURING ATTENTIONAL SET-SHIFTING TASKS

Jiachao Wang

A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews



2018

Full metadata for this item is available in  
St Andrews Research Repository  
at:

<http://research-repository.st-andrews.ac.uk/>

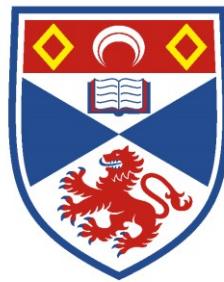
Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/14843>

This item is protected by original copyright

Bayesian analysis for quantification of individual rat  
and human behavioural patterns during  
attentional set-shifting tasks

Jiachao Wang



University of  
St Andrews

This thesis is submitted in partial fulfillment for the degree of PhD  
at the  
University of St Andrews

6<sup>th</sup> September 2017



**1. Candidate's declarations:**

I, Jiachao Wang, hereby certify that this thesis, which is approximately 42,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2014 and as a candidate for the degree of PhD in September 2014; the higher study for which this is a record was carried out in the University of St Andrews between 2014 and 2017.

I, Jiachao Wang, received assistance in the writing of this thesis in respect of grammar and syntax, which was provided by Doreen du Boulay, Sandra Piai, Katie White, and Dr Helen Sunderland.

Date: 31 August, 2017          Signature of candidate: Jiachao Wang

**2. Supervisor's declaration:**

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date    3 September 2017          signature of supervisor

**3. Permission for publication:**

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

a) No embargo on print copy

ELECTRONIC COPY

a) No embargo on electronic copy

Date: 6<sup>th</sup> September, 2017          signature of candidate          Jiachao Wang

signature of supervisor

## **Abstract**

Attentional set-shifting tasks, consisting of multiple stages of discrimination learning, have been widely used in animals and humans to investigate behavioural flexibility. However, there are several learning criteria (e.g., 6-correct-choice-in-a-row, or 10-out-of-12-correct) by which a subject might be judged to have learned a discrimination. Furthermore, the currently frequentist approach does not provide a detailed analysis of individual performance. In this PhD study, a large set of archival data of rats performing a 7-stage intra-dimensional/extra-dimensional (ID/ED) attentional set-shifting task was analysed, using a novel Bayesian analytical approach, to estimate each rat's learning processes over its trials within the task. The analysis showed that the Bayesian learning criterion may be an appropriate alternative to the frequentist *n*-correct-in-a-row criterion for studying performance. The individual analysis of rats' behaviour using the Bayesian model also suggested that the rats responded according to a number of irrelevant spatial and perceptual information sources before the correct stimulus-reward association was established. The efficacy of the Bayesian analysis of individual subjects' behaviour and the appropriateness of the Bayesian learning criterion were also supported by the analysis of simulated data in which the behavioural choices in the task were generated by known rules. Additionally, the efficacy was also supported by analysis of human behaviour during an analogous human 7-stage attentional set-shifting task, where participants' detailed learning processes were collected based on their trial-by-trial oral report. Further, an extended Bayesian approach, which considers the effects of feedback (correct *vs* incorrect) after each response in the task, can even help infer whether individual human participants have formed an attentional set, which is crucial when applying the set-shifting task to an evaluation of cognitive flexibility. Overall, this study demonstrates that the Bayesian approach can yield additional information not available to the conventional frequentist approach. Future work could include refining the rat Bayesian model and the development of an adaptive trial design.

## **Acknowledgement**

I would like to express my deep gratitude to my principal supervisor Dr Eric Bowman for his excellent instructions throughout the whole process of my PhD study. He is also a great role model to encourage me to keep on pursuing my academic career. I sincerely appreciate Dr David Tait, my secondary supervisor, for his detailed and prompt comments on my thesis drafts. I am really inspired by his expertise in neuroscience. I would also like to thank Prof. Verity Brown who provided the golden opportunity for me to pursue my PhD study. It would be impossible to finish my PhD research without her support in many ways.

The rat data used in Section 3 were collected by Dr Alexander Chase, Dr David Tait, and others from 2006 to 2009. I thank Ellen Bowman for efficiently transferring the rat data from paper-sheets to the electronic version. For the human data, Meggie Rix kindly volunteered to translate the participants' oral reports. In addition, I would like to thank my lab mates, Sonny, Shuang, Jenny, Ana, Casmira and Rudi for their kind help on PhD research.

I very much appreciate Dr Michael Oram for his essential advice on my study. I also received much warm encouragement and help during my PhD study from Dr Amanda Seed, Prof. Andrew Whiten, Dr Barbara Dritschel, Cheng Liu, Dr Gerry Quinn, Dr Ines Jentzsch, Janie Brooks, Dr Kate Cross, Dr Paula Miles, Zhemeng Wu and others. My sincere thanks also go to Dr Helen Sunderland, Doreen du Boulay, Sandra Piai, and Katie White who professionally proof-read my work. In addition, many thanks to my brothers and sisters in my Church for their kind sharing, advice and love.

I was honored to receive the St Leonard's College Scholarship from the University of St Andrews to financially support my PhD study over the past three years. Last but not the least, I would like to thank my family for their unwavering love. My mum flew to the UK especially to give me a warm hug, inspiring me in many ways.

## **Research Data Access Statement**

Research data underpinning this thesis are available at

<http://doi.org/10.17630/38f4d940-d16d-4250-a845-c0f43d524a2f>

## Contents

Abstract.....	i
Acknowledgement .....	ii
<b>Chapter I: Literature Review.....</b>	<b>1</b>
1 Introduction.....	1
1.1 Attentional set-shifting tasks .....	2
1.2 Significance of attentional set-shifting tasks .....	9
1.2.1 Human studies .....	10
1.2.2 Animal studies.....	11
1.3 Bayes' rule.....	22
1.4 Issues in task design and frequentist approach to data analysis .....	24
1.5 Utilising a simple Bayesian approach to analysing the 4ID task.....	27
<b>Chapter II: Bayesian Analysis of Rat Attentional Set-shifting.....</b>	<b>28</b>
2 Materials and Methods.....	28
2.1 Animals and apparatus .....	28
2.2 The 7-stage task.....	29
2.3 Hypothetical response patterns.....	31
2.4 Posterior probability of each hypothesis on a given trial .....	34
3 Results from the empirical rat research.....	47
3.1 p-value versus Bayesian estimate .....	47
3.2 Bayesian learning criterion is not always consistent with 6-in-a-row criterion .....	51
3.3 Bayesian analysis of individual rat data .....	53
3.4 Both lesioned and normal rats tried multiple spatial response patterns.....	55
3.5 Lesioned rats tried more spatial trials in ED stage than normal rats .....	57
3.6 Lesioned rats tried more spatial trials in reversal stages than normal rats .....	58
3.7 Association between length of spatial pattern and learning .....	59
3.8 First spatial response patterns at the ED are different between two rat groups.....	60
3.9 Overtraining reversal effect .....	61
4 Discussion on Bayesian analysis of rat data.....	64
4.1 Bayesian approach.....	64
4.2 Bayesian learning criterion.....	66
4.3 Bayesian analysis of individual rat's performance .....	67
4.4 Rats tried spatial response patterns in both control and lesioned groups .....	70



4.5	Lesioned rats tried spatial patterns for more trials than control rats.....	70
4.6	First spatial hypothesis in ED.....	73
4.7	Overtraining reversal effect.....	73
4.8	Summary .....	74
5	Bayesian analysis of simulated data.....	75
5.1	Individual analysis on simulated data.....	75
5.2	Bayesian analysis of spatial patterns in simulated data.....	82
5.3	Evaluation of Bayesian criterion in simulated data .....	85
5.4	Justification of the likelihood function.....	90
	<b>Chapter III: Bayesian Analysis of Human Set-shifting .....</b>	<b>93</b>
6	Method for human attentional set-shifting.....	93
6.1	Participants and apparatus .....	93
6.2	Materials and procedure .....	94
6.3	Human behavioural simple Bayesian model .....	98
6.4	Human latent probabilistic model and behavioural reward Bayesian model .....	100
6.5	Data screening .....	113
7	Results.....	114
7.1	Trials to criterion .....	114
7.2	Can Bayesian model accurately estimate which erroneous rules were used? .....	114
7.3	Individual estimate of set-formation before the ED stage.....	117
7.4	Latent probabilistic model can more accurately estimate participants' confidence .....	119
7.5	Bayesian learning criterion based on the latent probabilistic model performs better than the simple Bayesian model.....	121
7.6	Human latent probabilistic model better predicts whether participants are using correct response patterns to make choices.....	123
7.7	The human behavioural reward Bayesian model does not work well on rats .....	124
8	Discussion on Bayesian analysis of the human task .....	127
8.1	Human reward Bayesian model can accurately estimate erroneous rules used by participants .....	127
8.2	Human reward Bayesian model can estimate set formation for each individual.....	128
8.3	Human latent probabilistic model can more reliably estimate participants' confidence than behavioural simple Bayesian model.....	129
8.4	Human latent probabilistic model can better estimate participants' learning .....	129
8.5	Humans and rats may use different learning approaches .....	131

9	Human set-shifting task with Gabor patches.....	132
9.1	Trials to criterion .....	133
9.2	Can the human behavioural reward Bayesian model reliably estimate which erroneous rules were used?.....	133
9.3	Bayesian analysis can estimate individual set formation .....	135
9.4	Bayesian learning criterion based on the latent probabilistic model performs better...	136
9.5	Discussion .....	138
	<b>Chapter IV: Future Work and Conclusions.....</b>	<b>140</b>
10	Future work .....	140
10.1	Refinement of Bayesian model .....	140
10.2	Adaptive design.....	140
10.3	Bayesian learning criterion with fixed design .....	144
11	Conclusions .....	146
	References.....	149
	Appendices.....	159
	Appendix A: Robustness of rat model and human model.....	159

# **Chapter I: Literature Review**

## **1 Introduction**

Evolution has resulted in notable adaptability in behaviour in both humans and other animals, from which various forms of intelligence can be inferred (Mackintosh, 1974). One aspect of intelligence is the ability to cope flexibly with a mutable environment. For example, in daily life, people can easily follow and engage in varying topics when chatting with friends; when hearing a fire alarm, people can immediately stop what they are doing or thinking and escape from indoors; when a student takes a maths exam, the student can quickly shift across multiple methods and ideas in order to solve different questions. Such flexible switches or adjustments in thinking, attention or behaviours in response to changing environments or goals come from the mental ability sometimes called ‘cognitive flexibility’ (Scott, 1962; Dajani & Uddin, 2015).

While such adaptability is often overlooked in everyday life because of its pervasiveness, any deficit or impairment in cognitive flexibility can be devastating for the quality of life for those affected. People with cognitive flexibility deficits often perform poorly in demanding tasks which require shifts of attention or thinking from one aspect to another. For example, they may have difficulty in simultaneously doing two or more daily tasks like chatting and cooking, in updating their old beliefs or knowledge during learning, or in handling various emergencies at home or the workplace. Adaptability can be degraded in normal ageing and as a result of many neurological and psychiatric diseases, including Alzheimer’s disease, Parkinson’s disease, Huntington’s disease, schizophrenia, addiction, problematic gambling, obsessive compulsive disorder (OCD), autism, bipolar disorder, and anxiety disorder (Aloi et al., 2015; Bissonette & Powell, 2012; Bissonette et al., 2013; Fineberg et al., 2015; Floresco & Jentsch, 2011; Klanker et al., 2013; Van Eylen et al., 2015; Vazey & Aston-Jones, 2012; Young et al., 2009). Therefore, understanding the neural,

biochemical, and genetic mechanisms of cognitive flexibility will have significant scientific and social impacts.

### 1.1 Attentional set-shifting tasks

Various attentional set-shifting tasks have been developed to measure and assess putative cognitive flexibility. The most well-known task to infer cognitive flexibility is the Wisconsin Card Sorting Test (WCST) (Figure 1.1; Berg, 1948). In Berg's study, each participant was required to learn to match each card correctly to one of four reference cards based on the experimenter's feedback ('right' or 'wrong'), without being told what underlying category was being used. 60 cards each had a number of coloured shapes. One of three sorting categories was used: the same number of shapes, or the same colour of shape, or the same shape. After the subject correctly sorted five consecutive cards, the underlying category was changed, and the participant had to find the new category based on the experimenter's feedback, requiring a shift of his or her attention. In Berg's study, three cycles of category learning were performed, with each cycle including three learning stages corresponding to the randomly ordered categories. Later variants of the task (e.g., Milner, 1963; see Brown & Tait, 2016 for review) use 128 cards, with two cycles of the three categories, and 10 consecutive correct as the learning criterion.

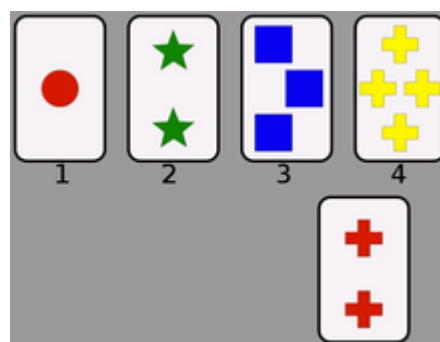


Figure 1.1: the Wisconsin Card Sorting Test. The participant must learn to match the card below to one of the four cards above, where the correct matching category is either based on colour, number of shapes, or the shape itself. Image: <https://commons.wikimedia.org/wiki/File:WisconsinCardSort.png>.

While the WCST has been widely used to infer human cognitive flexibility, performance in the WCST may be affected by other factors. For example, participants

may have worse WCST performance due to reduced sensitivity to experimenter feedback, or poor short-term memory (see discussion on human WCST performance in the 1<sup>st</sup> page of Tait et al., 2014). To isolate complicated, noisy cognitive components from those mediating behavioural flexibility, much simpler two-choice discrimination learning was adopted. Most of these tasks share a similar procedure, i.e., a series of two-choice discrimination learning stages. Across two learning stages, the reward-associated stimulus (or cue) is changed either within the same ‘dimension’ (called *intra-dimensional*, or ID) or to a different dimension (called *extra-dimensional*, or ED) (Eimas, 1966; Roberts et al., 1988; Settlage et al., 1956; Shepp & Schrier, 1969). In these ID/ED tasks, a ‘*dimension*’ implicitly refers to one type of feature of perceptual stimuli, i.e., one dimension refers to one perceptual property of the stimuli. For example, in the rat ID/ED task (Birrell and Brown, 2000), rats dig in one of two bowls and the stimuli in each bowl consist of a specific medium (e.g., ‘sawdust’) with a specific odour (e.g., ‘mint’). Therefore ‘odour’ would be one dimension and ‘digging medium’ would be the other dimension of the stimuli (see Table 1.1 for more examples). At the ID acquisition stage, subjects need to maintain their attention on the current reward-relevant dimension to learn about the correct/incorrect status of two novel stimuli from that dimension, while two novel stimuli from the other dimension remain irrelevant, e.g., from two odour stimuli ‘Cinnamon’ and ‘Ginger’ in the previous learning stage to another two odour stimuli ‘Sage’ and ‘Paprika’ in the new learning stage. In comparison, at the ED shift stage, subjects need to shift attention from a previously reward-relevant stimulus dimension to the previously irrelevant dimension, e.g., shift attention from ‘odour’ stimuli to ‘medium’ stimuli, also with ‘total change’ from old stimuli to novel stimuli in both dimensions (see Table 2.1 for a specific example of intra-dimensional and extra-dimensional shift stages). With the ‘total change’ from old stimuli to novel stimuli, partial reinforcement of any old stimuli from previous learning can be avoided in the new learning stages, such that the reinforcement effect for the specific (old) stimuli is eliminated but reinforcement effects of the perceptual dimension are retained.

When a particular dimension is repeatedly reinforced as reward-relevant, with others being irrelevant, a subject mentally forms an '*attentional set*', i.e., tendency or preference to attend to that particular dimension of the stimuli. Attentional set can be embodied as preparedness, based on prior experience, to attend to one aspect (or dimension) of multi-dimensional stimuli, which then facilitates learning about that aspect of those stimuli (Tait et al., 2014). Learning set, in a similar context, is preparedness, based on prior experience, to learn about the stimuli. Considering that attentional set is often used to influence discrimination learning in set-shifting tasks, attentional set could be considered as a specific form of learning set. Note that the formation of an attentional set can only be inferred from behavioural measurements, i.e., if the subject takes significantly more trials to reach the criterion in the ED shift stage than in the previous ID stage, this is viewed as indicative that the subject has formed an attentional set. If an attentional set is formed, in order to succeed at the ED stage, the subject needs to 'shift' attention from the previously formed attentional set to the previously irrelevant stimulus dimension, and hence the phrase '*attentional set-shifting*'. During attentional set-shifting, the previously formed attentional set often interferes with attentional shifting to the other dimension. When subjects are less behaviourally flexible, they take longer to shift to the other dimension. The resolution of the conflict between the previously relevant stimulus dimension and the newly-relevant stimulus dimension, as measured by the trials to the criterion in the ED stage, is a measure of behavioural flexibility.

In terms of ID and ED, each change in sorting category in the WCST requires an attentional shift, as in an ED stage of the ID/ED task. However, the WCST differs from the ID/ED task in some other aspects. For example, during each sorting category in WCST, all stimuli are present, with no stimuli being 'correct' or 'incorrect', and participants need to match a new card to one of four reference cards in each trial – unlike in the two-choice learning required as part of the ID/ED task. Also, one ID/ED stage just uses two possible stimuli for each of two dimensions, resulting in four possible stimuli combinations, while each WCST stage uses four possible stimuli for

each of three dimensions (colour, shape, and number), resulting in 64 possible cards. Therefore, while one ID/ED stage just requires participants to suppress one dimension of irrelevant stimuli, each WCST stage requires participants to suppress two irrelevant dimensions. In addition, each ID/ED stage uses totally new stimuli, while WCST may repeatedly use the cards over stages. Last but not least, the nature of the suppression is different at some stages. For instance, in the ID/ED task there is reversal learning, which probably require much more effort to suppress the attention to the unrewarded stimulus than in the WCST. In tasks measuring behavioural flexibility, both ID stage(s) and ED stages are typically included. As one of such attentional set-shifting tasks, the ID/ED task in the Cambridge Neuropsychological Automated Testing Battery (CANTAB) has been widely used in both humans and non-human primates such as marmosets (Dias et al., 1996; Roberts et al., 1988; Jazbec et al., 2007). The task consists of a series of two-choice visual discriminations between either different line segments or abstract shapes on a touch screen, including a simple discrimination (SD; where stimuli differ along only one dimension), compound discrimination (CD; where a second but reward-irrelevant dimension is included), ID acquisition stage, ED shift stage, and reversal learning stages (REV; where the previously rewarded stimulus becomes incorrect and the other previously incorrect stimulus of the same dimension becomes rewarded). A typical series of discriminations in the CANTAB ID/ED task has the order of simple discrimination– reversal learning – compound discrimination – reversal learning – ID – reversal learning – ED – reversal learning (Roberts et al., 1988; Table 1.1).

The CANTAB ID/ED task has also been adapted for rats (Birrell & Brown, 2000). However, rats seem reluctant to discriminate visual cues and often take hundreds of trials to finish training and testing over multiple days, which is time-consuming and limits the applications of the ID/ED task particularly in measuring acute effects of experimental manipulations (Tait et al., 2014). To avoid such issues, the rat version of the ID/ED task was developed to exploit the natural foraging behaviour of rats and uses digging bowls filled with scented media (Birrell & Brown, 2000). In the task,

which typically has seven stages (SD-CD-REV1-ID-REV2-ED-REV3; the ‘7-stage task’, Table 1.1), rats learn to discriminate bowls and dig in one of two paired bowls for a food reward based on one relevant dimension of stimuli in each learning stage. The reward-relevant stimulus dimension could be digging medium in the bowls or odour applied to the digging medium. Since rats naturally use odour and tactile cues for foraging, they typically learn to discriminate between two bowls of both dimensions within about 90 minutes.

While the 7-stage task has been widely used in rats, the early reversal stages may not always be sufficient to encourage set-formation, e.g., in rats with orbital prefrontal cortex (OFC) lesions (Chase et al., 2012). To enhance set-formation, a modified version of the 7-stage task has been used which excludes the reversal stages and includes three more ID stages (‘4ID’ in Table 1.1; Bissonette et al., 2008; Chase et al., 2012; Clarke et al., 2005). A cost between the last ID stage and the ED stage (i.e., taking more trials to learn the ED stage) indicates successful set-formation. Since ED deficits can be inferred to reflect attentional set-shifting deficits only when an attentional set can be concluded to have been formed, the 4ID task is more sensitive to set-formation-impaired subjects compared to the standard 7-stage task and can allow measures of set-shifting ability in instances where set is not formed in the 7-stage task. This is important because a majority of behavioural flexibility studies on rats involve evaluating ED shift ability in both control and experimental conditions, e.g., following manipulation of brain function through lesions, neural inactivation, or administration of neuroactive substances.

Similar to the CANTAB ID/ED task for primates, an automated touch-screen task using visual cues was developed for mice, but without reversal stages (Brigman et al., 2005) (‘Touch-screen’ in Table 1.1). This task requires mice to nose-poke the appropriate stimulus on the screen in order to receive reward from a food well under the screen, and wider varieties of visual stimuli can be easily displayed on the screen. Brigman et al. used this task for a between-subjects study, (i.e., one group of mice did



SD-CD-ID, and the other group did SD-CD-ED), and reported no difference between ID and ED performance. This suggests that mice may not have formed attentional set before the ID or ED stage, probably because there was insufficient learning experience, (i.e., only the SD and CD stage), before the ID or ED stage. On the other hand, since this task does not require body movement, it can be extended and applied to motor-deficit mice and to mice with genetic or pharmacologic manipulations (Bissonette & Powell, 2012; Graybeal et al., 2014; Mar et al., 2013; Marquardt et al., 2017).

Different from the CANTAB ID/ED task and its variants, strategy- or rule-shifting tasks have been developed with fewer learning stages and without novel stimuli across stages. One is the plus-maze (or cross-maze) task (Block et al, 2007; Floresco et al., 2006; Ragozzino et al., 2002). In this task, rats first learn to turn right (or left) in a plus maze for food reward without considering the starting arm and which arm contains a visual cue, namely *response discrimination*. Then, rats learn to turn to the arm which contains the visual cue for food reward, without considering whether it is a left or right turn, namely *visual-cue discrimination* (Ragozzino, 2002; Ragozzino et al., 2002). In this task, the rule needs to be shifted from spatial dimension to perceptual dimension, which is different from the 7-stage task where attentional set needs to be shifted between two different perceptual dimensions. The recently developed water T-maze task for mice is similar to the cross-maze task, but is performed in a water T-maze with additional reversal learning stages (Brooks et al., 2012; Table 1.1).

Another similar strategy-shifting task is the automated lever-pressing task, operated in a chamber (Brady & Floresco, 2015; Floresco et al, 2008). In this task, rats first learn to press the lever for food reward where the light above is illuminated, without considering whether the lever with the illuminated light is on the left or right. Then rats learn to press the right lever (i.e., ‘position response’ learning) without considering whether the light is illuminated above the right or left lever (Floresco et al,

2008). The automated lever-pressing task reduced the experimenter's manual effort during the rat's learning and dramatically improved throughput, while the 7-stage task requires an experimenter to manually configure bowls with food filled in one bowl for each learning trial. Compared to the 7-stage task, the (plus or T) maze task and the lever-pressing task did not change stimuli across learning stages.

More recently, one completely automated nose-poking ID/ED task has been developed for rodents (Scheggia & Papaleo, 2016), where three stimuli from three dimensions are provided, including light stimuli, texture, and odour. Since no manual work is required for each learning trial and the number and order of learning stages can be flexibly set with accompanying software, this new ID/ED task can be used for a large throughput of attentional set-shifting study on mice. However, the stimuli in the task are not compound (i.e., not mixed together in each chamber), meaning that there is a need for attentional reorienting spatially whilst ED-shifting. Any 'cost' seen at the ED might derive from the need to reorient attention rather than (or as well as) 'shift' from one perceptual dimension to the other.

While the animal set-shifting or strategy-shifting tasks were adapted from relevant human tasks such that the stimuli and configurations are more appropriate for these species to perform the tasks, they share the same purpose as that of human tasks, i.e., measuring one or more aspects of cognitive flexibility by shifting attention over two or more discrimination learning stages. Animal studies of cognitive flexibility with these shifting tasks have been widely adopted because it is difficult to directly investigate human brains due to ethical issues. As shown in the following (Section 1.2.2), the animal shifting tasks have helped researchers deeply understand the mechanism of cognitive flexibility from multiple levels.

Table 1.1: attentional set-shifting tasks and strategy-shifting tasks. CS: card sorting; Rev: reversal learning; ID1-ID4: four different ID shifts; VC: visual-cue discrimination; TR: turn-direction response learning; PR: position response learning; rTR: turn-direction response learning after water T-maze is rotated; ‘-’: from one learning stage to the subsequent learning stage; ‘m/n’: m correct response in n consecutive trials.

<i>Tasks</i>	<i>Stimulus dimension</i>	<i>Procedure</i>	<i>Learning criteria</i>	<i>p value given criteria</i>	<i>Species</i>
WCST (Berg, 1948)	Colour, number, shape	9 consecutive CS	5/5	0.031	Human
CANTAB ID/ED (Roberts et al., 1988)	Shape, line	SD-Rev-CD-Rev-ID-Rev-ED-Rev	Rev: 18/20 Others: 54/60	$1.8 \times 10^{-4}$ $< 10^{-10}$	Human, marmoset
7-stage (Birrel & Brown, 2000)	Odour, digging medium	SD-CD-Rev-ID-Rev-ED-Rev	6/6	0.016	Rats
4ID (Bissonette et al., 2008; Chase et al., 2012)	Odour, digging medium	SD-CD-ID1-ID2-ID3-ID4-ED	6/6	0.016	Mice Rats
Automated ID/ED task (Scheggia & Papaleo, 2016)	Odour, texture, light	e.g. SD-CD-Rev-ID1-Rev-ID2-Rev-ED-Rev	8/10	0.044	Mice
Plus-maze (or cross maze) (Floresco et al., 2006; Ragozzino et al., 2002)	Turn direction, visual cue	TR-VC	10/10	0.001	Rats
Lever-pressing (Floresco et al., 2008)	Position, visual cue	VC-PR	8/8	0.004	Rats
Touch-screen (Brigman et al., 2005)	Shape, line	SD-CD-ID/ED	16/20	0.005	Mice
Water T-maze (Brooks et al., 2012)	Turn direction, visual cue	TR-Rev-VC-Rev-TR-rTR	10/12	0.016	Mice

## 1.2 Significance of attentional set-shifting tasks

These attentional set-shifting tasks and their variations play a key role in human and animal behavioural studies for understanding various health issues (e.g., mental disorders and ageing) at different levels (including brain regions, neurotransmission, and genes) and for developing relevant pharmacological therapies. If the performance of subjects from the experimental group differs from that of the control group in a relevant set-shifting task, e.g., experimental subjects take longer to learn the ED stage than control subjects, it would suggest that the experimental factor somehow influences the set-shifting ability.

### 1.2.1 Human studies

Set-shifting tasks have played a key role in relating various human disorders and relevant deficits in cognitive flexibility. In human studies, the WCST and CANTAB ID/ED tasks have shown that apparent cognitive rigidity is a symptom in many neurological and psychiatric disorders with frontal lobe dysfunction or disconnection (Bissonette et al., 2013; for review see Brown & Tait, 2016; Klanker et al., 2013). It has long been known that patients with damaged frontal lobes showed impaired performance in the WCST (Berg, 1948). In the WCST, schizophrenia patients can learn the categories to sort the cards, but have great difficulty adapting to changes to the category (Egan et al., 2001; Prentice et al., 2008). Schizophrenic patients also showed impaired reversal learning in the CANTAB ID/ED task (Elliott et al. 1995; Leeson et al., 2009; Waltz & Gold, 2007). With the WCST, patients with depression (Borkowska & Rybakowski 2001), mild Alzheimer's disease (Nagahama et al. 2003; Perry et al. 2000), Parkinson's disease (PD) (Lange et al., 2016), and autism spectrum disorder (Westwood et al., 2016) showed perseveration errors when card-sorting categories were changed. Lange et al. (2016) recently showed that perseveration errors in PD patients may result from the neural impairments of multiple executive processes relevant to set-shifting process. With the CANTAB ID/ED task, impaired ED shifting ability in patients with Huntington's (HD) and PD can be easily detected even at the early stages of the diseases (Lawrence et al., 1996; Owen et al., 1992; Robbins, 2007). Further analysis with a novel stimulus dimension replacing the previously reward-relevant dimension showed that the ED shift deficit in HD comes from perseveration to the previously relevant dimension (Lawrence et al., 1999). In addition, patients with eating disorders showed an ED shift deficit, which also appeared in their first degree relatives, suggesting that genetic factors may play a role in eating disorder and certain (unknown) cognitive process might affect set-shifting and contribute to eating disorders (Treasure & Schmidt, 2013). An ED shift deficit was also found in obsessive compulsive disorder (OCD) (Chamberlain et al., 2006). All these disorders have been shown to be associated with frontostriatal circuit

disruption (Chudasama & Robbins, 2006). With the sensitivity of attentional set-shifting tasks to various neurodegenerative disorders, the CANTAB ID/ED task and the WCST have become standard clinical practice to evaluate cognitive flexibility in neuropsychological assessments (Vazey & Aston-Jones, 2012).

### **1.2.2 Animal studies**

Studies of non-human mammals using attentional set-shifting tasks confirm that the frontal cortex is crucial for cognitive flexibility, as it is in humans. Across species, the ED shift has been shown to be selectively impaired in humans with frontal cortex damage (Manes et al., 2002; Owen et al., 1991), in non-human primates with lateral PFC damage (Dias et al., 1996, 1997), and in rodents with medial frontal cortex damage (Birrell & Brown, 2000; Bissonette et al., 2008; Ghods-Sharifi et al., 2008; Hamilton & Brigman, 2015; see left region in Figure 1.2). While the anatomy of frontal cortex varies across species, the consistent findings across species on ED shift may indicate that medial frontal cortex in rodents is analogous or even homologous to primate lateral prefrontal cortex (Brown & Bowman, 2002; Brown & Tait, 2016).

Different from ED shifting, reversal learning performance was shown to be selectively impaired in marmosets (Dias et al., 1996, 1997) and rodents (Bissonette et al., 2008; Boulougouris et al., 2007; Brigman et al., 2013; Hamilton & Brigman, 2015; Izquierdo et al., 2013; McAlonan & Brown, 2003) with orbitofrontal cortex (OFC, top left region in Figure 1.2) damage. However, with the 4ID task, Chase et al. (2012) found that rats with OFC lesions also showed impaired learning in the series of IDs, which suggests that OFC is not only crucial for reversal learning, but also plays a key role in set-formation (Chase et al., 2012). Interestingly, in OFC-lesioned rats performing the 4ID task, ED shifting was found to be impaired as well, indicating that OFC affects both reversal and set-shifting abilities either directly or as a mediator. On the other hand, a recent study shows that excitotoxic lesion of OFC sub-regions (including Walker's areas 11, 13 and 14) does not affect reversal learning in macaque

monkeys (Rudebeck & Murray, 2011; Rudebeck et al., 2013), but damage to the white matter laterally adjacent to OFC caused reversal deficits in macaque monkeys (Rudebeck & Murray, 2011; Rudebeck et al., 2013; Chau et al., 2015), suggesting the involvement of nearby white matter in reversal learning impairment which was thought to be caused only by OFC lesion (at least for macaque monkeys). Further study is required to clarify the mixed findings on the role of OFC.

Reversal learning and ID performance may be affected not just by the OFC and its surrounding area. A recent study showed that rats with lesions of the ventral midline thalamus (i.e., nucleus reuniens) showed significant impairment in the first reversal learning stage and the subsequent ID stage in the 7-stage set-shifting task (Linley et al., 2016). The nucleus reuniens is densely and reciprocally connected with hippocampus and the mPFC, and therefore the lesioned nucleus reuniens may directly cause the dysfunction of its connected areas, which play crucial roles in set-shifting (Linley et al., 2016). In addition, while lesion of either mPFC or hippocampus in rats does not impair reversal learning, the lesion of the two regions together impaired spatial reversal learning significantly with a strategy-shifting task (Mala et al., 2015). This suggests that multiple non-OFC brain areas could cooperatively affect spatial reversal learning, although the cooperative effect on perceptual discrimination reversal learning remains to be explored with set-shifting tasks.

Furthermore, with the 4ID task, Lindgren et al. (2013) found rats with dorsomedial striatum (DMS, dorsal and ventral caudate in Figure 1.2) lesions had a similar learning performance between ED and previous ID stages, which suggests the attentional set was not formed before the ED stage in these DMS-lesioned rats. This indicates DMS may have a role in set-formation. Set-formation may also be affected by anterior cingulate cortex (ACC, top left region in Figure 1.2), because rats with ACC lesions also showed no ID/ED difference (Ng et al. 2007). Also, the anterior and posterior cingulate cortex is found to selectively degrade ID performance (Ng et al. 2007; Kim et al., 2016), suggesting that normal functioning of the cingulate cortex is

crucial for intra-dimensional acquisitions. In addition, rats with subthalamic nucleus (STN) lesions are also impaired in set-formation (Xia et al., unpublished; Tait et al., 2017). All these findings indicate that set-formation may be a complex process involving the functions of multiple interrelated brain regions, such that impairments might arise from dysfunction at multiple points in the neural circuit.

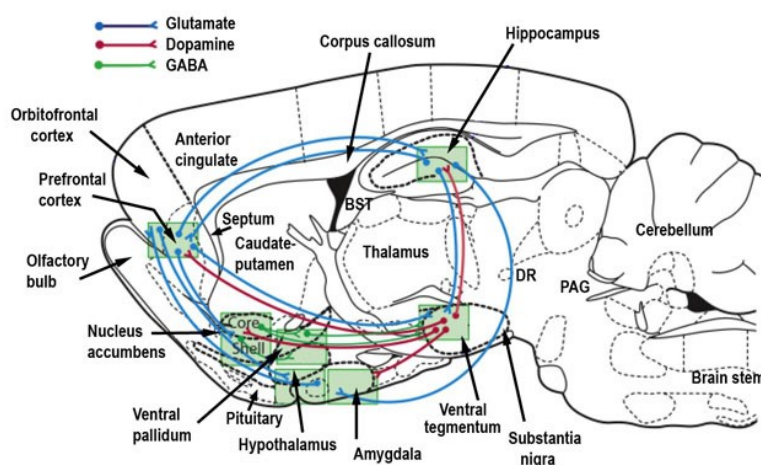


Figure 1.2: a sagittal view of rat brain anatomy, with three neurotransmitter pathways. The dopamine pathway (red) starts from ventral tegmentum (lower-right green box); glutaminergic pathways (blue) link prefrontal cortex (far left green box), hippocampus (upper-right green box), and basal ganglia (mid-bottom green boxes). GABAergic inhibitory neuron pathway (green curve) links ventral tegmentum to ventral pallidum and to nucleus accumbens (middle-left green boxes). BST = bed nucleus of the stria terminalis. DR = dorsal raphe nucleus; PAG = periaqueductal grey area. This figure is to roughly show the brain structure of rats and neural pathways across relevant brain regions, rather than to show the effect of various manipulations. From <http://sites.sinauer.com/animalcommunication2e/chapter10.04.html>.

Still with the 4ID task, Wright et al. (2015) found that rats with anterior thalamic damage showed impaired ID performance compared to control subjects. Since the thalamus-lesioned rats also finished the ED stage more rapidly than the previous ID stage, the damage of anterior thalamus is not disrupting the set-formation; otherwise, rats would have a similar performance in ED and ID. This study revealed that the anterior thalamus is crucial in directing attention to task-relevant stimuli particularly for ID learning, but at the expense of ED shift performance (Wright et al., 2015).

Similarly, the shifting between different dimensions or rules probably also requires proper interactions among multiple brain regions. For example, the pharmacological blockade of communication between the PFC and the perirhinal cortex (PER) impaired rats' learning performance when shifting reward-associated rules in an object-location paired association task (Hernandez et al., 2017). Considering that PER is an intermediate area which is reciprocally connected to mPFC and hippocampus, and the finding that hippocampus activity is associated with the inhibition of an incorrect response (Lee & Byeon, 2014), this study suggests that the mPFC-PER-hippocampus circuit is important for rule- or set-shifting ability (Hernandez et al., 2017). What's more, a recent study with a touch-screen set-shifting task in mice showed that the damage of the cerebellum (in terms of global loss of Purkinje cells in cerebellum due to mutation) would impair both reversal learning and ED shift (Dickson et al., 2017), supporting that normal functioning of cognitive flexibility is supported by an even larger brain circuit.

From the above brief review of animal studies of cognitive flexibility, it is clear that various shifting tasks have been playing a key role in exploring the neural mechanism of cognitive flexibility. In particular, different parts of the frontal cortex and striatum are involved in the learning at different stages of the task.

Besides helping to explore brain region functions, these attentional set-shifting tasks have also helped in the development of animal models of schizophrenia. The cognitive deficits, particularly an ED shifting impairment, in human patients with schizophrenia is associated with a reduction in prefrontal dopaminergic neurotransmission (Harrison, 1999). Therefore, rat models of schizophrenia can be induced by administering drugs that directly target the prefrontal dopaminergic system. For example, the administration of the drug phencyclidine (PCP) or ketamine has been shown to impair ED shift performance in rats by indirectly inducing dopaminergic dysfunction in the frontal cortex, leading to a potential animal model of schizophrenia (Jentsch & Roth, 1999; Nikiforuk et al., 2010; see Tait et al., 2014 for a



review). A similar finding was reported by developmental administration of ketamine in mice (Jeevakumar et al., 2015). With this schizophrenia model, Nikiforuk et al. (2016) found that the positive allosteric modulators (PAMs) of the alpha-7-nicotinic acetylcholine receptors can effectively reverse the ED deficit in rats, providing a potential treatment of schizophrenia with the PAMs. Recently, the phosphodiesterase 10A selective inhibitor TAK-063 was also found to remove the ED deficit and other cognitive dysfunctions associated with schizophrenia in rodent model, suggesting that TAK-063 may potentially ameliorate cognitive deficits of schizophrenia (Shiraishi et al., 2016).

As another example, aged rats showed impairment in set-shifting ability compared to young rats just as in aged humans, suggesting that the aged rat model can be used to investigate possible ageing causes and potential pharmacy to slow down the ageing process or mediate ageing-related cognitive dysfunction (Beas et al., 2013). Beas et al. (2017) recently showed that lower expression of GABA receptors in mPFC is strongly correlated with worse strategy-shifting performance in aged rats, and GABA receptor agonist baclofen enhanced their set-shifting performance, suggesting that GABA receptors may play a necessary role in affecting ageing-related cognitive flexibility ability. Finally, brief and repeated exposure to cocaine leads to reversal learning impairments, suggesting a rat model of the impacts of drug use (Izquierdo et al., 2010; Seu & Jentsch, 2009; see Izquierdo & Jentsch, 2012 for a review), and physical exercise in rats improved both reversal and ED performance, supporting the positive effect of sport on cognitive flexibility (Brockett et al., 2015). Such animal models provide convenient ways to develop and validate novel pharmacological agents and therapies for relevant disorders and diseases with cognitive flexibility symptoms, or even to improve health.

Since dysfunctions or lesions in the frontostriatal circuit have been shown associated with impaired cognitive flexibility, it is expected that any interruption of neurotransmitter pathways in the frontostriatal circuit would also influence cognitive

flexibility abilities. Together with animal models, shifting tasks have further helped understand cognitive flexibility at the neurochemical level, particularly the effect of neurotransmitters and receptors on various disorders.

### Dopamine (DA)

The dopamine pathway (red curves in Figure 1.2) in the frontal cortex and basal ganglia plays a key role in reward-motivated learning behaviour. Therefore, disruption of DA activity in the frontostriatal circuit is expected to influence discrimination learning in attentional set-shifting tasks (Glimcher, 2011). Studies with a maze-based task found that the blockade of either D<sub>1</sub> or D<sub>2</sub> receptors in rodent mPFC induces severe perseveration errors in strategy-shifting, while D<sub>1</sub> and D<sub>2</sub> agonists have no effect on strategy-shifting (Ragozzino, 2002; Floresco et al., 2006). Unlike D<sub>1</sub> and D<sub>2</sub>, stimulation of D<sub>4</sub> receptors in PFC results in more perseveration errors in strategy-shifting, whereas the blockade of D<sub>4</sub> improves strategy-shifting performance. Furthermore, depletion of PFC DA impairs ID (Crofts et al., 2001) but improves ED (Roberts et al., 1994) in marmosets, which suggests that PFC DA depletion impairs attentional set-formation. These findings suggest that various DA receptors in PFC may need to cooperate well for normal shifting performance (Floresco, 2013).

Besides the effect of PFC DA on ID and ED shifting performance, DA in the striatum is also involved in cognitive flexibility. In particular, studies found that DA transmission depletion in DMS selectively impairs reversal learning in both primates (Clarke et al., 2011) and rats (O'Neil & Brown, 2007). In comparison, in the ventral striatum, the D<sub>2</sub> receptor agonist may lead to more perseveration errors in both reversal learning and shifting (Haluk & Floresco, 2009; Yawata et al., 2012), while the D<sub>1</sub> receptor blockade may lead to more strategy-acquisition errors (i.e., taking more trials to acquire the reward-relevant strategy) in mice measured by number of 12-trial-long learning sessions (Yawata et al., 2012). Subsequently, a recent study in rats found that the synaptic DA level in the ventromedial striatum increases in response to unexpected reward in a lever-based reversal learning paradigm, and such increase in

DA level was observed only when rats successfully learned the reversal (Klanker et al., 2015), further supporting that striatal DA is involved in the adaptation of previously learned behaviour. All together, these findings support that striatal DA receptors are crucial in disengagement of previous learning and establishment of new learning (Yawata et al., 2012).

In vivo measurement of extracellular DA levels shows that DA levels in nucleus accumbens (part of ventral striatum, lower left region in Figure 1.2) increased when shifting rules on a T-maze-based strategy-shifting task, and DA levels in mPFC increased when both acquiring and shifting rules (Stefani & Moghaddam, 2006). Consistently, the role of DA in cognitive flexibility has also been observed in humans. For example, in patients with PD, administration of L-DOPA reversed the ED shift deficit in the CANTAB ID/ED task, while it impaired reversal learning performance (Cools, 2006; Kehagia et al., 2010).

## Serotonin

Unlike the DA effect in prefrontal-striatal circuit on set-shifting and reversal learning, serotonin (5-hydroxytryptamine; 5-HT) in the OFC only was found to affect reversal learning. Specifically, in marmosets, depletion of OFC 5-HT (but not DA) impaired reversal learning with significantly more perseveration errors (Clark et al., 2005, 2007), while 5-HT depletion in the medial caudate nucleus did not (Clarke et al., 2011). Consistently, with a touch screen reversal task, inactivation of the 5-HT transporter (resulting in increased 5-HT levels and, potentially, DA levels, for the 5-HT transporter can take up DA at least in rodents) improves reversal learning (particularly by decreasing perseveration errors) compared to controls in mice, suggesting that 5-HT transporter loss may increase sensitivity to negative feedback, and pharmacological treatments leading to the loss of the 5-HT transporter could reduce impairments in reversal learning (Brigman et al., 2010).

Interestingly, a recent study found depletion of 5-HT in primate OFC impaired

reversal learning when DA level in the putamen was also low, but did not impair reversal ability when putamen DA level was high. This suggests that the reversal learning may be affected by the balance between 5-HT levels in OFC and DA levels in the dorsal striatum (Groman et al., 2013).

### Norepinephrine (NE)

Attentional set-shifting tasks helped confirm that NE in mPFC plays a role in ED shifting. The norepinephrine (NE; also called Noradrenaline or NA) autoreceptor antagonist atipamezole (Lapiz & Morilak, 2006) or treatment of NE reuptake blocker (Lapiz et al., 2007), both of which increase NE neurotransmitter levels in the mPFC, were found to improve ED shift performance in the 7-stage task. Consistently, NE depletion (Tait et al. 2007) or NE deafferentation (McGaughy et al. 2008) in mPFC selectively cause ED shift impairment. Given that NE is mainly synthesised in the locus coeruleus and is projected to both mPFC and OFC, it is not surprising that optogenetic silencing of locus coeruleus activity impaired both ED shift and reversal learning performance, but did not impair ID and set-formation in an attentional set-shifting task on mice (Janitzky et al., 2015). The set-shifting deficit by NE deafferentation (McGaughy et al. 2008) can be reduced by increasing NE activity via blocking NE reuptake (Newman et al., 2008). Also, treatment with an NE reuptake inhibitor was found to be able to prevent or reverse an ED shift deficit originally caused by chronic stress (Bondi et al., 2008; Nikiforuk & Popik, 2011; Naegeli et al., 2013). Because NE neurons were found to be lost in early PD patients who also show impaired set-shifting ability (Vazey & Aston-Jones, 2012), and chronic stress is characteristic of depressive patients, the investigation of the NE effect on cognitive flexibility may help develop new therapies for early PD and stress in depression (Naegeli et al., 2013; Vazey & Aston-Jones, 2012).

### Acetylcholine (ACh)

Besides DA, 5-HT, and NE, acetylcholine (ACh) in DMS has also been confirmed as being associated with cognitive flexibility. An ACh antagonist infused into rat DMS

caused a reversal learning impairment (McCool et al., 2008), but there is no effect on rats' cognitive flexibility from cholinergic lesions in basal forebrain and PFC regions (McGaughy et al., 2008; Tait & Brown, 2008). Consistently, cognitive rigidity was found to be associated with a degraded ACh receptor M1 in rat DMS (Nieves-Martinez et al., 2012), and inactivation of cholinergic interneurons (which release ACh) in striatum caused more perseveration errors when shifting to a new strategy in rats (Aoki et al., 2015). Also, because ACh synthesis in striatum often degrades in aged rats (Das et al., 2001) and aged rats often show impaired reversal learning deficits (Brushfield et al., 2008; Schoenbaum et al., 2002), degraded ACh activity may underlie some of the cognitive deficits in ageing that are reflected in set-shifting performance. In this regard, increasing striatal ACh activity via AChE (acetylcholinesterase) inhibitor tacrine reduces ageing-related cognitive deficit (Tait et al., 2013), and the ACh receptor agonist nicotine improved both ID and ED shift performance in normal rats (Allison & Shoaib, 2013). All the evidence supports that ACh in DMS is crucial for normal function of cognitive flexibility, in particular for reversal learning.

### **N-methyl-D-aspartate (NMDA)**

Many studies using shifting tasks have shown that NMDA receptor antagonists injected into mPFC impair behavioural flexibility. The NMDA antagonists phencyclidine (PCP) and ketamine were found to impair reversal learning in rats (Abdul-Monim et al., 2006; Floresco et al., 2009; Idris et al., 2010), and PCP and another antagonist MK-801 (dizocilpine) were found to impair ED shift in rats (Rodefer et al., 2005; Stefani & Moghaddam, 2005). Blockade of the NMDA GluN2B receptor in PFC (Dalton et al., 2011), and knockout of NMDA GluN2A receptors (Marquardt et al., 2014), also impaired set-shifting, and cortex-wide deletion of GluN2B (Radke et al., 2015), or the local blockade of GluN2B in the lateral OFC (Thompson et al., 2015), caused reversal deficits in a touch-screen task. Because inactivation or blockade of NMDA receptors produces negative symptoms and cognitive deficits of schizophrenia, but DA agonists do not, investigation of the

effects of NMDA manipulations on behavioural flexibility in attentional set-shifting tasks using rodents may facilitate the exploration of pharmacological therapies for schizophrenia (Neil et al., 2010).

Recently, Jett et al. (2017) showed that blocking NMDA or AMPA ( $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptors in the rat mPFC, causing compromised glutamate neurotransmission, is associated with chronic stress-induced ED deficits. On the other hand, acute administration of NMDA receptor antagonist ketamine, which increases glutamate neurotransmission via AMPA receptors in mPFC, reversed chronic stress-induced ED deficits (Jett et al., 2015). These results suggest that investigation of the role of NMDA receptors on chronic stress may help explore therapy for stress-related psychiatric disorders like depression.

A recent interesting study showed that multiple neurotransmitters could cooperatively affect behavioural flexibility. While individual administration of a low-dose of either NMDA receptor antagonist or D<sub>1</sub> receptor antagonist did not impair strategy-shifting performance in a lever-press task, the administration of both antagonists together did impair task performance (Desai et al., 2017), suggesting that subtle abnormality in different types of neurotransmitters may act cooperatively to cause a deficit in behavioural flexibility.

The above brief summary clearly shows that multiple neurotransmitters in the frontostriatal circuit influence cognitive flexibility from different aspects, either independently or cooperatively, with the help of shifting tasks. The effects of various neurotransmitters on cognitive flexibility have even been confirmed at the genetic level via studies of neurotransmitter-related genes, as seen below.

### Genetic factors

Shifting tasks used with transgenic mice have helped confirm the effects of various neurotransmitters on behavioural flexibility (Klanker et al., 2013). For example,

DARPP gene (dopamine- and cAMP-regulated neuronal phosphoprotein) is related to D<sub>1</sub> receptor activation, and DARPP-32 knockout mice showed a reversal learning deficit (Heyser et al., 2000). This supports that D<sub>1</sub> receptor activation is necessary in reversal learning. Researchers also found that mice with D<sub>2</sub> receptor (gene: *Drd2*) knockout showed perseveration errors in simple reversal learning (Kruzich & Grandy, 2004; Kruzich et al., 2006), while mice with D<sub>3</sub> receptor knockout showed enhanced performance in compound reversal learning (Glickstein et al., 2005). Interestingly, D<sub>2</sub> receptor knockout did not affect mice's ID and ED shift performance (DeSteno & Schmauss, 2009), suggesting that D<sub>2</sub> receptors may selectively regulate reversal learning. However, it is not clear whether the effect of genes on cognitive flexibility is direct or indirect (e.g., via increased neurotransmitter levels), and since receptor knockout affected the whole brain, it is not clear which brain region of the receptors causes the impairment (Klanker et al., 2013).

For the effect of GABA-related genes, studies found that mice with mutant *Met* or *Plaur* genes had fewer GABAergic interneurons in OFC and striatum and showed reversal learning impairments (Bissonette et al., 2010, 2015; Martins et al., 2011). The reversal learning deficit in mutant *Plaur* mice can be restored after postnatal supplementation of the HGF gene (Bissonette et al., 2010). Because all the three genes (*Met*, HGF, and *Plaur*) associated with GABAergic interneurons are found associated with schizophrenia (Torrey et al., 2005; Bissonette & Powell, 2012), such transgenic studies in mice may speed up the exploration of gene therapies for associated disorders like schizophrenia. A recent transgenic study on mice has indicated already that abnormality in GABAergic interneuron-driven Gamma oscillations in the PFC may play a key role in causing cognitive inflexibility in schizophrenia and provides a potential therapeutic strategy for the disorder (Cho et al., 2015).

Overall, extensive study findings on cognitive flexibility from multiple levels, as briefly demonstrated above, have clearly shown the significance of attentional set-

shifting tasks. Because of the importance of set-shifting tasks, any development of new methodology in analysing set-shifting task performance, and refinement of set-shifting tasks, would benefit the study of cognitive flexibility. For this purpose, in this PhD project, a new Bayesian approach has been developed. To make it easier to understand our Bayesian approach, I will first introduce the basics of the Bayesian approach in the following section, and will describe our Bayesian approach in the methodology section later.

### 1.3 Bayes' rule

The fundamental component of the Bayesian method, i.e., Bayes' rule (Gelman et al., 2003), is introduced here, considering its crucial role in the whole project. Bayes' rule provides a way to update one's belief after receiving new evidence. Mathematically, Bayes' rule combines currently observed data with a previously formed belief to estimate how likely *each* specific hypothesis  $h$  is true:

$$P(H = h | E) = \frac{P(E | H = h) \cdot P(H = h)}{P(E)} \quad (1)$$

The above formula consists of four probability functions  $P(\cdot)$  with two variables  $H$  and  $E$ :

- a)  $H$  is a variable called *hypothesis*, and  $h$  is a specific hypothesis from the set of possible hypotheses. For example, *nationality* can be assigned with a specific value 'UK', 'USA', or 'China', etc.
- b)  $E$  is a variable called *data* or *evidence*, consisting of a set of observed information. For example, the *data*  $E$  could include a student's appearance and English accent in order to estimate how likely the student's *nationality* is UK'.
- c)  $P(H = h | E)$  is the *posterior probability* function. It represents the posterior probability of hypothesis  $H$  being a specific value  $h$  given the data  $E$ , where '|' symbol represents 'given'. It measures how likely a specific hypothesis  $h$  is true *after* (therefore *posterior*) the new evidence  $E$  is observed. The



posterior probability is what people or researchers want to know. How can we obtain the posterior probability  $P(H = h | E)$ ? The above Bayes' rule tells us the posterior probability can be obtained by calculating the three functions on the right side of the formula.

- d)  $P(E | H = h)$  is a *likelihood* function, representing the likelihood of the hypothesis  $h$  given the evidence  $E$ . This function measures how likely the data  $E$  is observed *if* a specific hypothesis  $h$  is true. In other words, it measures the consistency between the observed data  $E$  and a specific hypothesis  $h$ . For example, if a student's nationality is 'UK', how likely is it that the student has an oriental face and standard British accent? In general, the likelihood function  $P(E | H = h)$  needs to be defined beforehand for any particular application (see Section 2.4). Once the likelihood function is designed, then given a specific hypothesis  $h$ , the likelihood function  $P(E | H = h)$  will give us a likelihood value for a particular observation  $E$ . Note that different observations of  $E$  may correspond to different likelihood values for the same hypothesis  $h$ .
- e)  $P(H = h)$  is the *prior probability* of a specific hypothesis  $h$  being true. It represents the degree of people's belief that a specific hypothesis  $h$  is true *before* (therefore 'prior') observing the evidence  $E$ . The prior information about the specific hypothesis  $h$  often comes from previous relevant experience.
- f)  $P(E)$  is called *marginal likelihood*, which is a normalisation factor to make sure that the sum of the posteriors  $P(H = h | E)$  over all the possible hypotheses  $h$ 's is 1.0.

Based on Bayes' rule, the posterior probability of every specific hypothesis  $h$  being true is estimated based on not only the consistency between the hypothesis  $h$  and the

new observed evidence  $E$ , but also the old prior belief that the hypothesis  $h$  is true. Bayes' theorem is about the strength of belief for each possible hypothesis, whereas the classical frequentist approach embedded in current learning criteria in set-shifting tasks only considers the confidence of a single hypothesis, i.e., the null hypothesis. Also, the prior belief information is not considered in the classical frequentist approach.

#### 1.4 Issues in task design and frequentist approach to data analysis

Attentional set-shifting tasks on animals have been extensively used to study behavioural flexibility. But how do we know whether animals have learned as a researcher intended during discrimination learning in the set-shifting tasks? In all animal-learning studies, we can only infer whether animals have learned the correct stimulus-reward association based on the pattern of the behavioural choices. The behavioural learning criterion adopted in the attentional set-shifting tasks is typically based on a frequentist approach using inferential statistics. For example, a criterion of 6-correct-choices-in-a-row given a null hypothesis of responding randomly ( $p = 0.05^6 = 0.0156$ ) was adopted to determine whether a rat has learned to find the reward-associated stimulus in each stage in the 7-stage task (Birrell & Brown, 2000). However, there are other criteria under which the rat might be judged to have learned a discrimination. Suppose a rat makes five correct choices, followed by one error, then again five correct choices followed by one error (Table 1.2). Although the chances of randomly choosing the correct bowls on 10 out of 12 trials is less than 2% (0.019), such a 10-out-of-12 correct choice would not satisfy the 6-correct-choice-in-a-row criterion. This raises two issues: (1) how big the 'window' over which performance is considered should be, for windows that are too large will not detect learning well, and windows that are too small are prone to statistical errors; and (2) must a subject (animal or human) perform perfectly to conclude that the subject has learned the contingencies in the task?

Table 1.2: an example of a rat's choice results in a learning stage.

Trial	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15-20
Choice	✗	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓✓✓✓✓✓✓

Another issue is that the 6-correct-choice-in-a-row criterion (or similar criteria; Brigman et al., 2005; Brooks et al., 2012; Floresco et al., 2006; see Table 1.1) takes a classical null-vs-alternative hypothesis inferential testing approach. The rationale behind this approach is that, if the null hypothesis is very unlikely true (e.g., less than 5% chance), then the null hypothesis is rejected. The problem is that once the null hypothesis is rejected, there are often multiple alternative hypotheses consistent with the data. Deciding which of these alternatives is correct is beyond the scope of null hypothesis testing. For the 6-correct-choices-in-a-row criterion, ‘randomly choosing a bowl’ is the null hypothesis, and ‘using the reward-associated stimulus for bowl choice’ is considered as the (only) alternative hypothesis. However, this is a potential error in statistical reasoning, because there are other alternative hypotheses about the response pattern of the rat that do not lead to correct responding or random responding. For example, a rat may always choose bowls on the same side. This means the traditional ‘null-versus-alternative’ hypothesis testing approach does not fully describe the behaviour. The consequence of this is that while the criterion may indicate learning, it discards other information about the animal’s choices. This could lead to either more trials being presented even after the animal has learned the discrimination (‘false negative’), or it could lead to finishing a stage even if the animal has not learned the discrimination (i.e., ‘false positive’). For example, if a rat correctly chooses bowls in six consecutive trials, with the six side choices ‘left-left-right-left-right-left’, the rat might choose bowls by alternatively changing side of choice in last five trials rather than by reward-associated information, leading to a ‘false positive’. In this case, the rat would probably be under-trained and need more training.

The problem of false positives increases as the number of trials increases. I have performed a computer simulation to estimate the likelihood of happening to respond correctly in six consecutive trials when virtual rats are actually responding randomly (Figure 1.3). The simulation result shows that the likelihood of false positives increases quickly with more trials, indicating that false positives may often happen

when using 6-in-a-row criterion to judge rats' learning, particularly when rats find it hard to learn. This becomes disproportionately problematic for studies of animals exhibiting impaired behavioural flexibility; in which case, false positives would likely happen when animals take longer to learn, thereby *reducing* the apparent magnitude of the impairment.

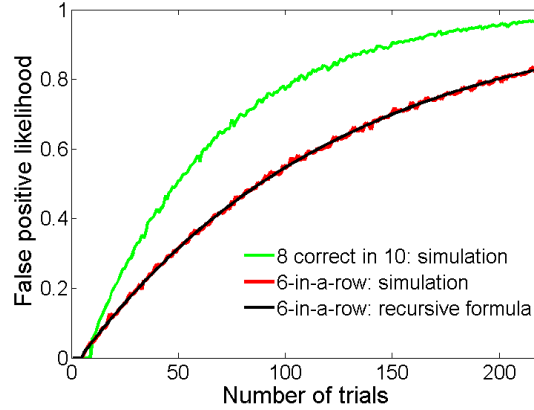


Figure 1.3: the likelihood of false positive increases with more trials with two different learning criteria. With the 6-in-a-row criterion (red curve), to estimate the false positive likelihood for a specific trial number  $k$ , I generated 5000 sequences of  $k$  binary (i.e., ‘correct’/‘incorrect’) values, with each binary value randomly generated. A sequence is considered ‘false positive’ if there exist six consecutive correct values anywhere in the sequence. The false positive likelihood is the ratio between the number of ‘false positive’ sequences and the total sequence number 5000. A similar simulation process was performed for the 8-correct-in-10 criterion (green curve). The simulation result with 6-in-row criterion (red curve) is confirmed by a recursive formula (black curve) recently described by Fazekas et al. (2010).

Moreover, the traditional hypothesis-testing approach can only help us decide when animals have learned, it does not allow us to determine which pattern of possible responding, which could be relevant to perceptual stimuli or spatial locations, predominates in each learning stage for each rat. Before establishing the correct stimulus-reward association, a rat may have tried other non-random but reward-irrelevant spatial patterns or stimulus characteristics on which to base a bowl choice (see Section 2.3 for detail). Knowing these details from each rat would help researchers understand more deeply the course of learning when animals solve discrimination learning problems, and explore the difference in learning processes between individuals and groups (e.g., the effects of different lesions or drug treatments). The approach based on null hypothesis  $p$ -values cannot provide such

detailed information.

### **1.5 Utilising a simple Bayesian approach to analysing the 4ID task**

I have implemented a simple Bayesian approach to avoid those issues mentioned above, by estimating the probability of each hypothetical pattern of behavioural choice at each learning trial. The Bayesian approach is not trying to model rats' decision processes from the viewpoint of a rat learning the task, but trying to determine how close the rats' responding matches a particular hypothetical pattern, as observed from the viewpoint of the experimenter.

In pilot work (unpublished undergraduate dissertation), the initial exploration of the Bayesian approach to analysing the archival data of an attentional set-shifting task (the 4ID task) in rats indicated that learning in some instances was stronger or weaker as judged from posterior probabilities of Bayesian analysis than the  $p$ -value indicated. Specifically, rats continued to be trained in some stages even when the Bayesian probability of the correct choice pattern was large (e.g.,  $>0.95$ ), while training stopped on other occasions even though the Bayesian posterior probabilities associated with the correct choice pattern were modest (e.g.,  $\sim 0.6$ ). This initial exploration also showed that the degree of learning may affect a rat's attentional set-shifting performance in the ED stage. To avoid a varying degree of learning in rats, I proposed replacing the traditional  $n$ -correct-choice-in-a-row criterion by the Bayesian learning criterion, i.e., the rat should finish a learning stage when the Bayesian (posterior) probability of the correct choice pattern becomes larger than a threshold (e.g., 0.95).

Since the frequentist  $p$ -value and Bayes' posterior probability can lead to different conclusions about when the rats have learned, it becomes necessary to explore the relationship between  $p$ -values and Bayes' posterior probability. Also, not all the observed information about rats' learning behaviour was considered in the initial Bayesian model. Including all the observed information would more likely result in a better mathematical model in accurately estimating the posterior probability of each

hypothetical pattern. In this thesis, I will extend the Bayesian model and further explore its applications in analysing attentional set-shifting task data.

## **Chapter II: Bayesian Analysis of Rat Attentional Set-shifting**

### **2 Materials and Methods**

#### **2.1 Animals and apparatus**

Forty-seven normal control rats with sham lesion surgery and forty-six medial prefrontal cortex (mPFC)-lesioned rats (Lister hooded; Harlan, UK) were used to perform the 7-stage task. Rats were maintained on a moderately restricted diet (15–20 g of food for each rat per day). On the day before each rat performed the task, each rat was exposed to the food-baited bowls and then was trained to complete two simple discrimination stages to a criterion of six consecutive correct trials. On the next day, the rat completed all seven stages of the task by satisfying the 6-correct-in-a-row criterion. In each learning stage, each rat dug one of two bowls in each trial to search for the food reward (Figure 2.1). Each bowl was distinguishable from the other in the trial by having a different digging medium inside the bowl and/or a different odour emanating from the digging medium. In each trial the configuration of the bowls and the rat's behavioural response was recorded (Figure 2.1), including which bowl was chosen by the rat, whether the chosen bowl contained a food reward, whether the rat dug after encountering one bowl (recorded as  $1^{st}$ ), or went to investigate the other bowl before digging (recorded as  $2^{nd}$ ), and the time spent on the trial. Based on the trial configuration generation rules, the reward-relevant stimulus would not appear on the same side for more than three times in a row. The archival data were collected from years 2006 to 2009 in the lab of Professor Verity Brown<sup>1</sup> (Tait et al., 2009). The data were originally collected and analysed for multiple attentional set-shifting studies. This is the first time that the Bayesian analysis has been applied to pooled data across

---

<sup>1</sup> The data were collected by Dr David Tait, Dr Alexander Chase, Sarah Hersman, Sarah Dennis, Robert Johns, Tamlyn Watermeyer, Anne Bremicker, Francesca Hand, Louise MacLellan, and Stephanie Hunter

multiple cohorts of rats. The reader is referred to Tait et al. (2009) for details about the apparatus and the testing method. All the data collection was carried out in accordance with the Animals (Scientific Procedures) Act 1986, under the Project licences granted to Prof. Verity J. Brown, with Project License (PPL) number 60/3138, 60/3837, and 60/4459, approved by the University of St Andrews Animal Welfare and Ethics Committee and the UK Home Office.

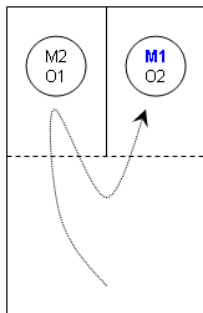


Figure 2.1: an example trial from the 7-stage task. In the apparatus, there are two chambers (upper half of the box). Within each chamber, there is one bowl (circle) containing one particular medium and one particular odour, e.g., the left bowl contains medium M2 and odour O1. The food is in only one of the two bowls and associated with either one particular medium or odour. Here, the food is associated with medium M1 (in bold blue). The curve with the arrow represents the rat's behaviour at this learning trial, i.e., the rat first approached the left bowl, but did not dig in it and then went to dig in the right bowl.

## 2.2 The 7-stage task

Briefly, after training on two simple two-choice discriminations in which stimuli differed along only one perceptual dimension, each rat was tested in a single day in a series of seven two-choice discrimination stages (Table 2.1). The first stage was a simple discrimination (SD) either between two odours (as an example in Table 2.1) or between two digging media. Rats learned which of the two stimuli was associated with a food reward. In the second stage, an irrelevant stimulus dimension was added to form a compound discrimination (CD), with the reward-associated stimulus unchanged compared to the previous SD stage. The third stage was reversal learning (REV1), where the reward-relevant dimension and all the stimuli remained unchanged compared to the CD stage. However, the reward-associated stimulus was reversed within the perceptual dimension (Table 2.1). Once the reversal learning was finished, another compound learning stage (ID) followed, in which the stimuli in both

dimensions were novel compared to previous stages, although the reward-relevant stimulus dimension was unchanged compared to the previous three stages. This compound stage was then followed by a second reversal learning stage (REV2). Through the sequence of stages from SD to REV2, the experiment was designed to reinforce the rat's attention to one dimension of stimulus (e.g., 'odour' in Table 2.1) in order to obtain the food reward. In the stage subsequent to REV2, the previously rewarded dimension ('odour') became irrelevant, and rats had to learn to attend to another dimension (e.g., 'medium' in Table 2.1) to obtain reward. In order to succeed in this extra-dimensional ('ED') discrimination stage, the rat in general had to shift its attention from one dimension ('odour') to the other dimension ('medium'), and therefore, rats generally required more trials to reach the learning criterion. After the ED stage, the task ended with another reversal (REV3). In the 7-stage task, the assignment order of stimuli pairs in each stage and the shift of dimension (from odour to medium, or vice versa) are counter-balanced insofar as possible.

Table 2.1: the 7-stage task. For each trial, bowls can be discriminated either between two odours or between two digging media. At each trial, the correct, reward-associated stimulus (in bold) is paired with one stimulus from the reward-irrelevant dimension, and pseudo-randomly assigned into either the left or the right bowl. M1-M6: six distinctive medium stimuli; O1-O6: six odour stimuli.

Discrimination learning stage	Dimensions		Stimulus combinations	
	Reward-associated	Reward-irrelevant	Rewarded combination	Un-rewarded combination
Simple (SD)	Odour	None	<b>O1</b>	O2
Compound (CD)	Odour	Medium	<b>O1</b> , M1	O2, M2
			<b>O1</b> , M2	O2, M1
Reversal (REV1)	Odour	Medium	<b>O2</b> , M1	O1, M2
			<b>O2</b> , M2	O1, M1
Intra-dimensional (ID)	Odour	Medium	<b>O3</b> , M3	O4, M4
			<b>O3</b> , M4	O4, M3
Reversal (REV2)	Odour	Medium	<b>O4</b> , M3	O3, M4
			<b>O4</b> , M4	O3, M3
Extra-dimensional shift (ED)	Medium	Odour	<b>M5</b> , O5	M6, O6
			<b>M5</b> , O6	M6, O5
Reversal (REV3)	Medium	Odour	<b>M6</b> , O5	M5, O6
			<b>M6</b> , O6	M5, O5



### 2.3 Hypothetical response patterns

In each discrimination learning stage of the set-shifting task, only one stimulus (e.g., a specific odour or medium) is associated with the food reward. A rat must learn to find the association between this stimulus and the food reward based on the feedback (i.e., either found food or not found food) from its choice of bowl on each trial. However, before the rats have learned the correct stimulus, there is no reason to believe that their choice of bowl is random. Rather, a rat might use various response patterns based on the stimulus configuration. There are at least two classes of external information about the bowls that the rat might use: *spatial* and *perceptual*. If a rat chooses a bowl based on the spatial locations of bowls rather than the stimuli in the bowls, then the rat would be using a *spatial response pattern* for bowl choice. Instead, if the rat chooses a bowl based on the stimuli (e.g., digging medium or odour) in the bowls, the rat would be using a *perceptual response pattern* for bowl choice. The plausible spatial and perceptual *response patterns* for bowl choice in the 7-stage task are listed below.

#### *Spatial response patterns:*

- Spatial alternation: if choosing the left (or right) side at the last trial, choose the right (or left) side for the current trial.
- Spatial perseveration: return to the same location as the previous trial, no matter whether the location is on the left or on the right. This would capture a side bias (left or right).
- Win-stay: if rewarded at the last trial, choose the same location for the current trial; otherwise choose the alternative location.
- Win-shift: if rewarded at the last trial, choose the alternative location for the current trial; otherwise, choose the same location.

*Perceptual response patterns:* there are four different stimuli combined across the two bowls, assuming no other perceptual characteristics are used given that the bowls are standardised: two odours (*O1* and *O2*) and two media (*M1* and *M2*). A

rat may choose a bowl based on one of the following perceptual response patterns:

- M1: choose the bowl which contains M1.
- M2: choose the bowl which contains M2.
- O1: choose the bowl which contains O1.
- O2: choose the bowl which contains O2.

Note that different perceptual stimuli may be used in different stages; therefore perceptual response patterns will be specific to particular stages. In comparison, the above four spatial response patterns may appear in any stage. Also note that here, only the simple response patterns were considered. More complex response patterns (e.g., the combination of O1 and M2) were not included, although they could be added if necessary. In addition, while the literature used the term *rules* to describe the above *response patterns*, we believe that *response pattern* is a more appropriate term, because (1) we can never know for sure what rules rats used, (2) even whether rats actually used any rules during learning, and (3) rats might make a selection to exclude a possibility even though they had a tentative rule in mind (e.g., a rat chose cumin maybe because it wanted to confirm it was incorrect rather than thinking cumin was correct).

Because we cannot know for sure which response pattern was actually used at each trial for each rat, we can only estimate the probability that the rat's choices match a given response pattern. More formally, we apply Bayes' rule to estimate the posterior probability of the following eight hypotheses regarding the possible response patterns at each trial:

- Hypothesis 1 ( $h_1$ ), or *spatial alternation* hypothesis: the rat uses spatial alternation response pattern to choose bowls.
- Hypothesis 2 ( $h_2$ ), or *spatial perseveration* hypothesis: the rat uses spatial perseveration response pattern to choose bowls.

- Hypothesis 3 ( $h_3$ ), or *spatial win-stay* hypothesis: the rat uses spatial win-stay response pattern to choose bowls.
- Hypothesis 4 ( $h_4$ ), or *spatial win-shift* hypothesis: the rat uses spatial win-shift response pattern to choose bowls.
- Hypothesis 5 ( $h_5$ ), or *M1* hypothesis: the rat uses perceptual M1 response pattern to choose bowls.
- Hypothesis 6 ( $h_6$ ), or *M2* hypothesis: the rat uses perceptual M2 response pattern to choose bowls.
- Hypothesis 7 ( $h_7$ ), or *O1* hypothesis: the rat uses perceptual O1 response pattern to choose bowls.
- Hypothesis 8 ( $h_8$ ), or *O2* hypothesis: the rat uses perceptual O2 response pattern to choose bowls.

Note that in the SD stage, only two (rather than four) stimuli appear in the bowls. Therefore, for the SD stage, there are in total only six hypotheses. On the other hand, more hypotheses would be generated if more complex response patterns were considered. However, Bayesian estimate of these eight hypotheses already provides much richer information compared to the hypothesis testing approach, which evaluates only one hypothesis linked to the random responding associated with the null hypothesis. Also, the eight hypotheses are used as a proof of concept to establish the utility of Bayesian analysis applied to learning. Other hypotheses could be added to the Bayesian analysis with limited effort if necessary. In addition, we assume that the eight hypotheses are disjointed (mutually exclusive of each other). This assumption is reasonable because the rat cannot simultaneously use two or more response patterns to make choice in a trial. In other words, considering one hypothesis being true would be exclusive of all the other seven hypotheses being true. I would like to clarify that such independence assumption describes relationship between the hypotheses in the same trial. Such assumption does not exclude the potential

relationships between hypotheses across trials. For example, if rats would try spatial response patterns first and then stimulus-based response patterns, the priors of stimulus-based hypotheses at current trial might be affected by how likely the rat tries both the stimulus-based response pattern and spatial response patterns from the previous trial. We did not explore this possibility because (1) we were exploring to begin with and this add yet another free variable to the analysis and (2) Occam’s razor – until we have concrete data that suggests this on balance we should assume a simpler mechanism. From the observation of Bayesian analysis results on multiple rats’ data, we found no evidence for such spatial-stimulus circle of patterns. Note as well that there are other potential dependences between the hypotheses across trials, for instance the rats always went with odour first and then medium. All such potential cross-trial dependence relationship does not contradict with the within-trial independence relationship between hypotheses.

## **2.4 Posterior probability of each hypothesis on a given trial**

It is not possible to know with certainty the basis on which a rat chooses bowls in the task, but it is possible to determine the degree to which the observed rat’s learning data across trials is consistent with each of the response pattern hypotheses. The observed data in each trial includes (1) the two perceptual characteristics (one odour and one medium) in the chosen bowl, (2) the rat’s chosen location, and (3) whether the rat dug after encountering only one, or both, bowls.

To estimate the probability of each of the eight hypotheses, we use the observed data from both the current trial and all previous trials (either within the stage or across stages, depending on which stage the current trial is in). Let  $E$  and  $E^*$  respectively denote the new observed data from the current trial and the old data from all previous trials, respectively. Then, all the observed data from the first trial to the current trial within a learning stage will be the combined observation  $\{E, E^*\}$ . To estimate the

probability of a specific hypothesis  $h_i$  (where  $h_i$  can be any of the listed hypotheses  $h_1$  to  $h_8$ ) after we observe the rat's data  $\{E, E^*\}$ , Bayes' rule tells us

$$P(H = h_i | \{E, E^*\}) = \frac{P(E | \{H = h_i, E^*\}) \cdot P(H = h_i | E^*)}{P(E | E^*)} \quad (2.1)$$

which is a simple extension of Equation (1.1) (Section 1.3) with the addition of the observation of previous trials,  $E^*$ . Note that including  $E^*$  in Equation (2.1) is the only difference to Equation (1.1).  $H$  is the hypothesis variable, the value of which could be any of the eight specific hypotheses. The comma ',' in the likelihood  $P(E | \{H = h_i, E^*\})$  means 'and'; therefore  $P(E | \{H = h_i, E^*\})$  represents the probability of  $E$  given  $H = h_i$  and  $E^*$ .

Equation (2.1) tells us that to estimate the posterior probability  $P(H = h_i | \{E, E^*\})$  of a specific hypothesis  $h_i$  in the current trial (i.e., left side of Equation 2.1), we need to compute the likelihood  $P(E | \{H = h_i, E^*\})$ , set the prior probability of the specific hypothesis  $h_i$ ,  $P(H = h_i | E^*)$ , and compute the marginal likelihood  $P(E | E^*)$ . Intuitively, the posterior probability  $P(H = h_i | \{E, E^*\})$  measures how likely a rat used the response pattern linked to the hypothesis  $h_i$  to make choices, after the digging behaviour at both current and previous trials was observed. The likelihood  $P(E | \{H = h_i, E^*\})$  measures how likely it is that we would observe the data  $E$  if a rat used the response pattern linked to hypothesis  $h_i$  to dig a bowl. The prior  $P(H = h_i | E^*)$  represents the researcher's prior confidence that the rat would use the response pattern linked to  $h_i$  to choose a bowl before the rat made a choice at the current trial. The marginal likelihood  $P(E | E^*)$  is simply a normalisation factor. In the following, we will introduce how to respectively compute the likelihood of spatial

hypotheses and perceptual hypotheses, and then how to compute the prior probability and the marginal likelihood.

## **Likelihood of spatial hypotheses**

The likelihood  $P(E | \{H = h_i, E^*\})$  measures how likely it is that we will observe the data  $E$  at the current trial if the rat used the specific response pattern corresponding to hypothesis  $h_i$  to dig in a bowl in the current trial. The observed data  $E^*$  appears in the likelihood because observed information from the previous trial would be involved in computing the likelihood of spatial hypotheses (e.g., win-stay; see below for details). Two types of information from the observed data in the current trial contribute to the likelihood computation.

### **Considering encountering one or both bowls before bowl choice**

One part of observation,  $E_1$ , is whether the rat dug after encountering one, or both, bowls. The former is more compatible with a spatial pattern of responding whereas the latter is more compatible with a pattern of responding based on the perceptual characteristics of the bowl. For example, if the rat's choice is based on the spatial perseveration and it had dug in the left bowl on the previous trial, then the rat would choose the same location (the left bowl) in the current trial. By contrast, if the rat approached either bowl and then moves to the alternative location, it is unlikely that the rat is using spatial information to guide its choice, because the rat would dig the first bowl it encounters if it was using a spatial response pattern (as discussed above). So, if we only consider the observed data about whether the rat dug in the first (i.e.,  $E_1 = 1^{st}$ ) or the second bowl ( $E_1 = 2^{nd}$ ), the likelihood of any spatial hypothesis  $h_i$  (either  $h_1$ ,  $h_2$ ,  $h_3$ , or  $h_4$ ) could be estimated by

$$P(E_1 | \{H = h_i, E^*\}) = \begin{cases} \alpha, & \text{if } E_1 = 1^{st} \\ 1 - \alpha, & \text{if } E_1 = 2^{nd} \end{cases} \quad (2.2)$$

In Equation (2.2), the parameter  $\alpha$  represents the high likelihood that a rat will choose to dig in whichever bowl it encounters first if it is using a spatial response pattern; therefore,  $\alpha$  should be set at a value close to 100% (i.e., 1.0). Nevertheless, allowing for the naturally inquisitive explore-exploit behaviour of these spontaneously foraging animals (Stephens, 2008), we set  $\alpha = 0.9$ , resulting in

$$P(E_1 | \{H = h_i, E^*\}) = \begin{cases} 0.9, & \text{if } E_1 = 1^{st} \\ 0.1, & \text{if } E_1 = 2^{nd} \end{cases}$$

Note that  $\alpha$  should not be set to 1.0, because otherwise the likelihood (Equation 2.2) could be 0.0 (i.e.,  $1 - \alpha$ ) for some spatial hypotheses in some trials, which in turn would cause the posterior probability of the corresponding hypotheses to be 0.0 (from Equation 2.1). Intuitively, the probability of any hypothesis should not be zero.

### **Considering location and stimuli in the chosen bowl**

So far, we have only considered the information regarding whether a rat chose to dig after encountering one, or both, bowls. Choosing the first bowl does not necessarily mean that the likelihood of a specific *spatial* hypothesis  $h_i$  is high. To finally determine the likelihood of a specific hypothesis  $h_i$ , in other words, to measure the consistency between the observed data and the specific response pattern corresponding to the hypothesis  $h_i$  in the current trial, we have to consider the other observed information,  $E_2$ , including the rat's chosen location and perceptual stimuli in the chosen bowls. Intuitively, if a rat dug in the left bowl at both the previous trial and the current trial, we would think the rat is more likely to use the spatial perseveration rather than spatial alternation response pattern to choose bowls, because the rat dug in the same-side bowl over the two trials. In this example, because the spatial perseveration response pattern is consistent with the rat's digging behaviour and the spatial alternation response pattern is not consistent with the rat's behaviour, the likelihood of the hypothesis corresponding to the spatial perseveration response pattern would be higher, and the likelihood of the hypothesis corresponding to the spatial alternation response pattern would be lower. Formally, the probability of

observing  $E_2$  will be higher if  $E_2$  is consistent with the hypothesis being considered, and vice versa. As a result, the above likelihood function (Equation 2.2) for any spatial hypothesis  $h_i$  can be further refined to

$$P(E_2, E_1 = 1^{st} | \{H = h_i, E^*\}) = \begin{cases} \alpha_1, & \text{if } E_2 \text{ consistent with } h_i \\ \alpha - \alpha_1, & \text{otherwise} \end{cases} \quad (2.3)$$

$$P(E_2, E_1 = 2^{nd} | \{H = h_i, E^*\}) = \begin{cases} \alpha_2, & \text{if } E_2 \text{ consistent with } h_i \\ 1 - \alpha - \alpha_2, & \text{otherwise} \end{cases} \quad (2.4)$$

The parameter  $\alpha_1$  represents the high likelihood of the *spatial* hypothesis  $h_i$  when  $h_i$  is consistent with the observation  $E_2$  at the current trial. Note that  $\alpha_1$  is less than  $\alpha$  (Equation 2.2), because otherwise zero or negative values ( $\alpha - \alpha_1$ ) could be assigned to the likelihood function (Equation 2.3), which would then cause zero or negative posterior probability (from Equation 2.1). Similarly, the parameter  $\alpha_2$  must be less than  $1 - \alpha$ . By default, we set  $\alpha_1 = 0.8$  and  $\alpha_2 = 0.05$ , resulting in

$$P(E_2, E_1 = 1^{st} | \{H = h_i, E^*\}) = \begin{cases} 0.8, & \text{if } E_2 \text{ consistent with } h_i \\ 0.1, & \text{otherwise} \end{cases}$$

$$P(E_2, E_1 = 2^{nd} | \{H = h_i, E^*\}) = \begin{cases} 0.05, & \text{if } E_2 \text{ consistent with } h_i \\ 0.05, & \text{otherwise} \end{cases}$$

Figure 2.2 shows an example of calculating the likelihood of each spatial (and perceptual) hypothesis at two different conditions: the rat digging in the bowl it first encountered ( $E_1 = 1^{st}$ ; Figure 2.2, lower left) or going to investigate the second bowl before making a response ( $E_1 = 2^{nd}$ ; Figure 2.2, lower right).



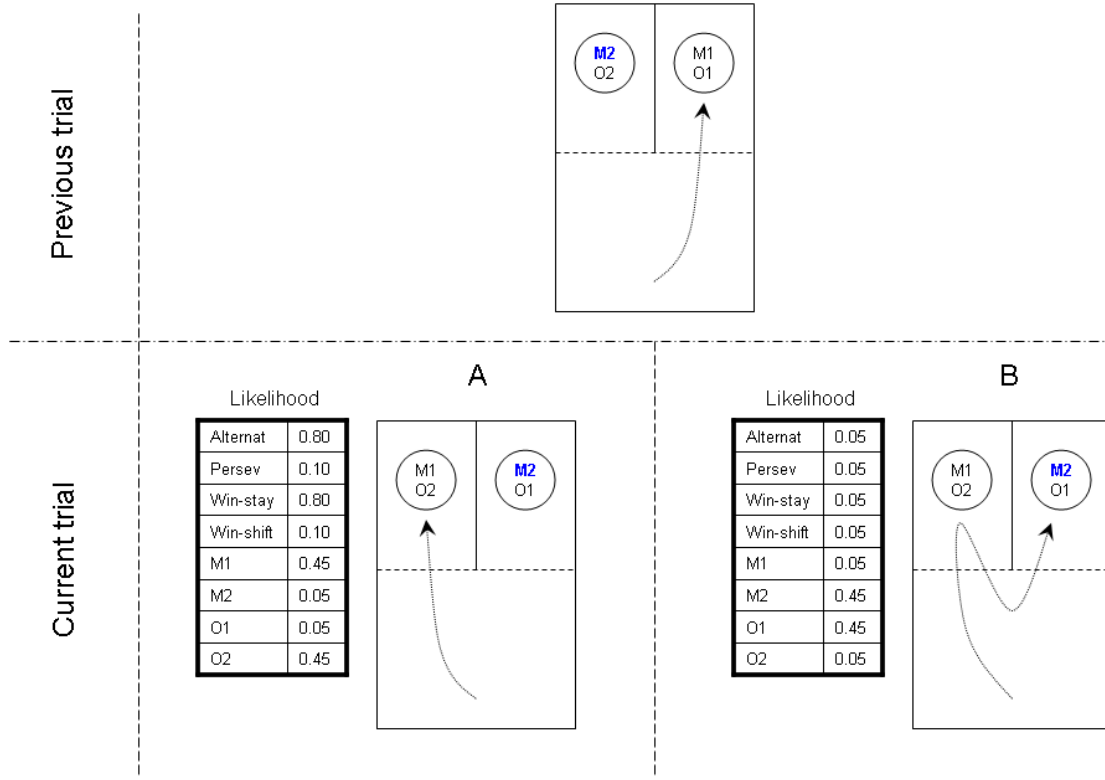


Figure 2.2: an example of calculating likelihood of each hypothesis based on the previous and current trials' observation. Suppose in the previous trial (upper figure), the rat directly went to dig in the right bowl which did not contain the food-associated  $M2$  (solid blue). Then, in the current trial, if the rat directly went to dig in the left bowl (lower left, subfigure A, curve with arrow), then  $E_1 = 1^{st}$ , and the likelihood of each spatial hypothesis (the lower-left table) can be calculated by Equation (2.3) and the likelihood of each perceptual hypothesis can be calculated by Equation (2.6). Instead, if the rat approached the left bowl but then went to dig the right bowl (lower right, subfigure B, curve with arrow), then  $E_1 = 2^{nd}$ , and the likelihood of each spatial hypothesis (collected in the lower-right table) can be calculated by Equation (2.4) and the likelihood of each perceptual hypothesis can be calculated by Equation (2.7). The calculation of the likelihood of spatial hypotheses depends on the rat's behaviour in both the current and the previous trial, while the calculation of the likelihood of perceptual hypotheses depends only on the rat's current trial behaviour. Note in the tables, 'Alternat' is the abbreviation of *Alternation*, and 'Perserv' for *Perseveration*.

## Likelihood of perceptual hypotheses

The above likelihood estimates are for *spatial* hypotheses. While the likelihood of any *perceptual* hypothesis will be computed the same way, the parameters in the likelihood function will be different for perceptual hypotheses.

Similar to the discussion on the likelihood of spatial hypotheses, we first look at how to estimate the likelihood of any *perceptual* hypothesis if we *only* consider the observed data about whether the rat dug after one, or both, bowls were encountered. Intuitively, because the medium and odour configuration at every trial is designed in a (pseudo-)random way, and therefore, the rat could not know which medium or odour is in which bowl, the probability of finding any specific medium or odour from the first bowl the rat encountered is equivalent to that from the second bowl. In other words, when the rat wants to choose a bowl containing a specific medium (or odour), the probability of finding the specific medium in the first bowl and then digging in the first bowl is equal to the probability of finding the specific medium in the second bowl and then digging in it. For example, if the rat is using the perceptual response pattern *M1* (corresponding perceptual hypothesis is  $h_5$ ) to choose which bowl to dig in, then the rat must approach one of the bowls, and the likelihood of encountering *M1* in the first bowl it approaches is only 50%. That means, the likelihood of digging in the first bowl it approaches will not be higher (or lower) than the likelihood of not digging in that bowl, but rather moving to the second bowl. Thus, if we *only* consider the observed data about whether the rat dug in the first bowl it encountered (i.e.,  $E_1 = 1^{st}$ ) or dug in one bowl after encountering the two bowls ( $E_1 = 2^{nd}$ ), the likelihood of any perceptual hypothesis  $h_i$  (either  $h_5$ ,  $h_6$ ,  $h_7$ , or  $h_8$ ) is

$$P(E_1 | \{H = h_i, E^*\}) = \begin{cases} \beta, & \text{if } E_1 = 1^{st} \\ 1 - \beta, & \text{if } E_1 = 2^{nd} \end{cases} \quad (2.5)$$

where  $\beta = 0.5$  represents the equal likelihood of choosing the first and the second bowl the rat encounters if it uses a perceptual response pattern to choose bowls, resulting in

$$P(E_1 | \{H = h_i, E^*\}) = \begin{cases} 0.5, & \text{if } E_1 = 1^{st} \\ 0.5, & \text{if } E_1 = 2^{nd} \end{cases}$$

Similar to the further division for spatial hypotheses (please see the rationale about how to further divide in the previous section), the above likelihood function for perceptual hypotheses can be further divided by considering the remaining observation  $E_2$  in the current trial,

$$P(E_2, E_1 = 1^{st} | \{H = h_i, E^*\}) = \begin{cases} \beta_1, & \text{if } E_2 \text{ consistent with } h_i \\ 0.5 - \beta_1, & \text{otherwise} \end{cases} \quad (2.6)$$

$$P(E_2, E_1 = 2^{nd} | \{H = h_i, E^*\}) = \begin{cases} \beta_2, & \text{if } E_2 \text{ consistent with } h_i \\ 0.5 - \beta_2, & \text{otherwise} \end{cases} \quad (2.7)$$

As for  $\alpha_1$  and  $\alpha_2$  in Equations (2.3) and (2.4), both  $\beta_1$  and  $\beta_2$  will be less than 0.5.

By default we set  $\beta_1 = \beta_2 = 0.45$ , resulting in

$$P(E_2, E_1 = 1^{st} | \{H = h_i, E^*\}) = \begin{cases} 0.45, & \text{if } E_2 \text{ consistent with } h_i \\ 0.05, & \text{otherwise} \end{cases}$$

$$P(E_2, E_1 = 2^{nd} | \{H = h_i, E^*\}) = \begin{cases} 0.45, & \text{if } E_2 \text{ consistent with } h_i \\ 0.05, & \text{otherwise} \end{cases}$$

Based on the above formulae, the likelihood of every specific hypothesis  $h_i$  in the current trial can be calculated by considering whether the rat chose the first bowl it encounters, and then by considering whether the rat's bowl choice is consistent with the specific response pattern corresponding to  $h_i$ . Figure 2.2 shows an example of calculating the likelihood of each spatial and perceptual hypothesis when the rat dug the bowl it firstly encountered ( $E_1 = 1^{st}$ ; Figure 2.2, lower left) or went to dig the second bowl ( $E_1 = 2^{nd}$ ; Figure 2.2, lower right).

## Prior probability

The prior probability  $P(H = h_i | E^*)$  represents the researcher's prior confidence that

the rat used the response pattern associated with  $h_i$  to choose a bowl before the rat started the current trial. Within a learning stage, if we estimate that the rat likely chooses a bowl on the last trial based on a specific response pattern associated with hypothesis  $h_i$ , then we would expect that the rat will also likely use the same response pattern to choose bowls in the current trial. Therefore, it is possible to use the posterior probability of the specific hypothesis  $h_i$  estimated from the last trial as the prior probability of the same hypothesis  $h_i$  for the current trial. In this way, the Bayesian approach combines the contribution of the new observed data in the current trial (reflected in likelihood function) with the contribution of observed data from previous trials (reflected in prior function).

However, special considerations are necessary for prior resetting in the first trial of each learning stage (Table 2.2). For the first trial of the first stage (SD) in the task, the rat has just begun the task and has not encountered any of the stimuli to be used in this stage. Therefore, we have no evidence to determine which response patterns were more likely used by rats; thus we assume all the possible hypotheses equiprobable with the prior probabilities for all hypotheses being set to 1/number of hypotheses (1/6).

For each of the other six stages, the stage begins when the rat has learned a particular perceptual feature that signals the location of the food reward in the previous learning stage. The stimuli in the new learning stage might be the *same* (e.g., in reversal stages), i.e., the same as those in the previous learning stage, or be *partial* (e.g., in CD stage), i.e., partially appear in the previous stage, or be *novel* (e.g., in ID and ED stages), i.e., novel exemplars of the same stimulus dimensions as the previous learning stage (Table 2.2, second column).

Table 2.2: prior setting for the first trial of each learning stage.

Stages	Consistency with previous stimuli	Prior in first trial
SD	N/A	Set as equal (i.e., 1/6) for all six hypotheses
CD	same + novel	Two new perceptual hypotheses' prior set with 1/8; All existing six hypotheses' priors maintained from SD
REV1	same	Priors of all eight hypotheses maintained from CD
ID	novel	Priors of all eight hypotheses reset as equal (i.e., 1/8)
REV2	same	Priors of all eight hypotheses maintained from ID
ED	novel	Priors of all eight hypotheses reset as equal (i.e., 1/8)
REV3	same	Priors of all eight hypotheses maintained from ED

For reversal stages, the prior of each hypothesis in the first trial of the new learning stage is the hypothesis' posterior probability in the last trial of the previous learning stage, because there is nothing to signal to the rat that conditions have changed. In other words, in this first trial, the rat would be expected to use the same hypothesis to choose bowls as it used on the previous (correct) trial.

For the CD stage, where the stimuli partly change with the addition of another (albeit irrelevant) dimension, considering the fact that the rat has learned the reward-associated stimulus, the rat would be expected to choose bowls with the same reward-associated perceptual response pattern. Therefore, the priors of the old stimuli in the first trial of the new learning stage are the same as their posterior probabilities in the last trial of the SD stage. For the priors of the two new perceptual hypotheses in this stage, since the rat prefers the reward-associated (old) response patterns, the initial priors of the two new hypotheses would be low, e.g., being 1/8 (considering the total number of hypotheses being 8). With similar reasoning, all spatial hypotheses' priors will be the same as their posteriors in the last trial of the previous stage.

Finally, for ID and ED stages, since all stimuli are novel compared to those in the previous stage, the stimulus of the rat's preference in the previous stage disappears and the rat would probably have no preference for any of the new stimuli. In other words, the rat would have equivalent preference for the stimuli. Therefore, the priors of all hypotheses (including spatial hypotheses) are reset to equiprobable initial priors

1/8 (if the total number of hypotheses is eight). In this condition, the priors of the spatial hypotheses are also reset to the initial prior, because rats may revert to using a spatial response pattern for making a choice when they do not know which, if any, perceptual stimulus might predict where the food is.

Another special consideration concerns the minimum prior. Intuitively, even if we have very low prior confidence that a rat will use a certain response pattern, the confidence will not be null, but rather will be at a low level. With that in mind, a lower bound is placed on all priors, such that the minimum value of each prior in any trial does not fall below a threshold, e.g., 0.001. In addition, note that the hypotheses' prior can be renormalised such that their sum is 1.0. However, this prior renormalisation is not necessary because of the renormalisation of the product of likelihood and prior for each hypothesis by the marginal likelihood (see below).

## Marginal likelihood

The other term in the Bayes' rule is the marginal likelihood  $P(E | E^*)$  in the current trial. As a normalisation factor, it is simply the sum of the product of likelihood and prior over all hypotheses. So the marginal likelihood  $P(E | E^*)$  can be directly calculated by summing the product  $P(E | \{H = h_i, E^*\}) \cdot P(H = h_i | E^*)$  over all hypotheses,

$$\begin{aligned} P(E | E^*) = & P(E | \{H = h_1, E^*\}) \cdot P(H = h_1 | E^*) + P(E | \{H = h_2, E^*\}) \cdot P(H = h_2 | E^*) \\ & + P(E | \{H = h_3, E^*\}) \cdot P(H = h_3 | E^*) + P(E | \{H = h_4, E^*\}) \cdot P(H = h_4 | E^*) \quad (2.8) \\ & + P(E | \{H = h_5, E^*\}) \cdot P(H = h_5 | E^*) + P(E | \{H = h_6, E^*\}) \cdot P(H = h_6 | E^*) \\ & + P(E | \{H = h_7, E^*\}) \cdot P(H = h_7 | E^*) + P(E | \{H = h_8, E^*\}) \cdot P(H = h_8 | E^*) \end{aligned}$$

## Posterior probability

Based on Bayes' rule (Equation 2.1), the posterior probability of each hypothesis,  $h_i$ ,

can be calculated as follows:

$$P(H = h_1 | \{E, E^*\}) = \frac{P(E | \{H = h_1, E^*\}) \cdot P(H = h_1 | E^*)}{P(E | E^*)}$$

$$P(H = h_2 | \{E, E^*\}) = \frac{P(E | \{H = h_2, E^*\}) \cdot P(H = h_2 | E^*)}{P(E | E^*)}$$

...

$$P(H = h_8 | \{E, E^*\}) = \frac{P(E | \{H = h_8, E^*\}) \cdot P(H = h_8 | E^*)}{P(E | E^*)}$$

Based on the formula of marginal likelihood and posteriors, the posterior probabilities of all the possible (i.e., eight) hypotheses will be summed to 1.0, as shown in the example below.

As an example, let's assume the likelihood of the eight hypotheses  $h_1$  to  $h_8$  are respectively 0.80, 0.10, 0.80, 0.10, 0.45, 0.05, 0.05, 0.45, and the prior of the eight hypotheses are respectively 0.02, 0.08, 0.17, 0.03, 0.56, 0.04, 0.07, 0.03. Then, the marginal likelihood can be computed by Equation (2.8),

$$\begin{aligned} P(E | E^*) &= 0.80 * 0.02 + 0.10 * 0.08 + 0.80 * 0.17 + 0.10 * 0.03 \\ &\quad + 0.45 * 0.56 + 0.05 * 0.04 + 0.05 * 0.07 + 0.45 * 0.03 \\ &= 0.434 \end{aligned}$$

And the posterior probabilities of the eight hypotheses based on the above formulae are, respectively

$$P(H = h_1 | \{E, E^*\}) = 0.80 * 0.02 / 0.434 \approx 0.037$$

$$P(H = h_2 | \{E, E^*\}) = 0.10 * 0.08 / 0.434 \approx 0.018$$

$$P(H = h_3 | \{E, E^*\}) = 0.80 * 0.17 / 0.434 \approx 0.313$$

$$P(H = h_4 | \{E, E^*\}) = 0.10 * 0.03 / 0.434 \approx 0.007$$

$$P(H = h_5 | \{E, E^*\}) = 0.45 * 0.56 / 0.434 \approx 0.581$$

$$P(H = h_6 | \{E, E^*\}) = 0.05 * 0.04 / 0.434 \approx 0.005$$

$$P(H = h_7 \mid \{E, E^*\}) = 0.05 * 0.07 / 0.434 \approx 0.008$$

$$P(H = h_8 \mid \{E, E^*\}) = 0.45 * 0.03 / 0.434 \approx 0.031$$

The sum of the eight posteriors is

$$0.037 + 0.018 + 0.313 + 0.007 + 0.581 + 0.005 + 0.008 + 0.031 = 1.0$$



### 3 Results from the empirical rat research

With Bayes' rule, every hypothesis' posterior probability was estimated for each trial in each learning stage of the 7-stage task for each rat. If a rat performed perfectly in the task, the posterior probability of the reward-associated perceptual hypothesis should soon be monotonically increased to a value close to 1.0 over trials in each learning stage, after making a few mistakes ('unlucky guesses') in choosing bowls at the beginning of the stage. The posterior probabilities of the remaining hypotheses should quickly decrease to small values close to 0.0 when the posterior probability of the rewarded hypothesis becomes large (close to 1.0). In the following, we used Bayesian probability information to explore Bayesian learning criterion, the detailed learning process of individual rats, and various spatial response patterns within and between rat groups. The computation of posterior probability of each hypothesis in each trial for each rat was implemented in MATLAB.

#### 3.1 p-value versus Bayesian estimate

The widely used 6-in-a-row learning criterion is based on the calculation of p-value for the null hypothesis that the rat is randomly choosing bowls. If the number of consecutively correct bowl choices is  $n$  since last wrong choice, then the p-value is  $p = 0.5^n$ . In this way, for every rat's data collected with the 7-stage task, a p-value can be calculated for every learning trial.

The Bayesian approach provides a possible alternative learning criterion to determine whether a rat has learned the reward-associated stimulus in each learning stage. More specifically, if the posterior probability of the reward-associated hypothesis is larger than a pre-defined high threshold (e.g., 0.95, meaning that we are 95% confident the rat is using the reward-associated stimulus to choose bowls) at certain trial, we would consider that the rat has learned the stimulus-reward association. The Bayesian learning criterion is based on the calculation of the Bayesian *probability of the*

*reward-associated hypothesis*. The Bayesian probability of the reward-associated hypothesis (called *b-value* in the following) can be calculated by the Bayesian approach for every learning trial.

The degree of correlation between the p-value (actually  $1-p$ ) and the b-value was computed to evaluate whether the Bayesian learning criterion might be an appropriate alternative to the 6-in-a-row criterion. If the correlation is too low, then it is unlikely that the Bayesian estimate will be useful, for the p-value is expected to be a good criterion for judging when the appropriate learning has occurred. If the correlation is too high (e.g.,  $>0.9$ ), then the b-value would be of little worth because it would not provide additional information compared to the p-value. The average correlation between the b-value and p-value across trials for all rats was found to be moderate and positive (mean correlation=0.57,  $n=93$ , range within one standard deviation [0.47, 0.66], corrected by Fisher's transform, Fisher 1921), suggesting that b-value broadly agrees with p-value but is partially independent of it.

However, it is possible that the psychological processes that operate in one (e.g., lesioned) group might make the correlation different to that of the other (e.g., control) group. Since we want the Bayesian technique to apply equally well to all groups of rats, it would be unsettling if the properties changed from group to group. For each rat, we obtained the correlation (called *P-B correlation*) between the p-value and the b-value over all the trials. As a result, we obtained 46 P-B correlation values for the 46 lesioned rats, and 47 P-B correlation values for the 47 control rats. Figure 3.1 (left) shows the variations of the P-B correlations for the lesion group and control group respectively. The P-B correlation for the control group (mean rank=52.57) is significantly higher than that for the lesion group (mean rank=41.30),  $U(97)=819$ ,  $Z=2.01$ ,  $p=0.044$ . This is potentially problematic, for it might indicate the properties of the Bayesian estimate varies by group.

For each rat, we not only calculated a P-B correlation, but also counted the total

number of trials the rat took to finish all the seven stages. Interestingly, higher P-B correlations often correspond to a smaller total number of trials, which more likely appears in control rats (Figure 3.1, right, filled dots). In contrast, lower P-B correlations corresponds to a higher total number of trials, which more likely appears in lesioned rats (Figure 3.1, right).

Thus, the apparent difference in P-B correlation between the two groups could be an emergent property of the number of trials the rats took to complete the task. A one-way ANOVA with group as a between-subjects factor and number of task trials as a covariate was conducted. There was a significant effect of the covariate ( $F(1,90)=40.2$ ,  $p<0.001$ , partial eta-squared=0.31), but not of group ( $F(1,90)=0.00$ ,  $p=0.98$ , partial eta-squared<0.001). Thus, the apparent difference between the two groups appears to reflect merely the number of trials taken to complete the task.

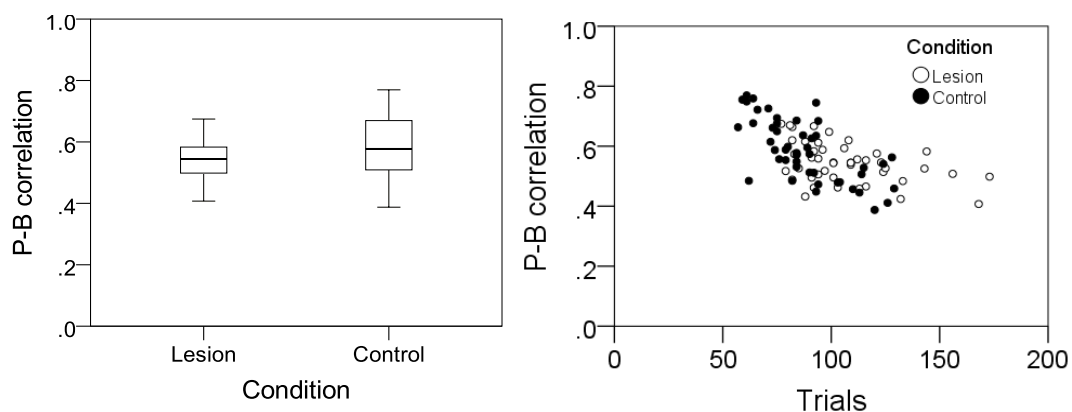


Figure 3.1: P-B correlations for both groups. Left: the P-B correlation for lesioned rats (left box-plot) is lower than that for control rats (right box-plot); Right: the scatter plot between the total number of trials required to complete all seven stages (x-axis) and the P-B correlation (y-axis) over all the control and lesioned rats. It should be noted that while the relationship between the total number of trials and the P-B correlation can be summarised with a linear function, the best fit would probably be a curvilinear function

But, a new question arises: why does the correlation between p-value and Bayesian estimate (i.e., P-B correlation) decrease with more task trials? We demonstrated above that the false positive likelihood of the 6-in-a-row criterion increases over trials (Pearson correlation coefficient = 0.974,  $p<0.01$ ,  $R^2=0.949$ ; Figure 1.3 in Section 1). The false positive likelihood (for corresponding total number of trials) and the P-B

correlation for each rat were found to be strongly correlated (Pearson correlation coefficient = -0.618,  $p < 0.01$ ,  $R^2 = 0.382$ ; Figure 3.2), as though the association between the p- and b-values breaks down as the possibility of false positives inflates over trials.

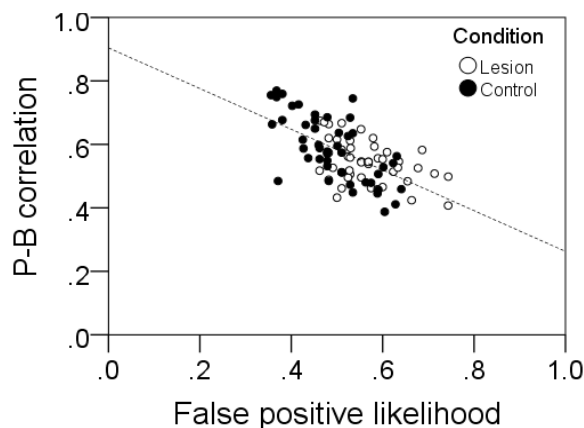


Figure 3.2: scatter plot of the correlation between p-value and b-value (i.e., P-B correlation) and the false positive likelihood corresponding to the total number of trials.

To further compare p-value and b-value, we also evaluated their performance in predicting whether a rat's next trial choice is correct based on the current trial's p- or b-value. For the prediction performance of p-value, an individual logistic regression was fitted to predict the next trial's choice correctness based on the current trial's p-value over all seven stages' learning trials for each rat. Similarly, a logistic regression was fitted using the b-value for each rat. Both p-value and b-value are significant predictors. This resulted in 93 individual regression models for both the p-value predictor and b-value predictor, respectively. B-value was the significant predictor ( $p < 0.05$ ) in 36 out of 93 rats, and p-value was the significant predictor ( $p < 0.05$ ) in 46 out of 93 rats. Therefore, b-value as significant predictor is not significantly different from p-value as significant predictor (chi-square(1)=2.18,  $p = 0.14$ , ns.). Furthermore, for both predictors, the residual of prediction for each trial was computed as the difference between predicted value (between 0 and 1) and the actual outcome (with '1' for correct choice and '0' for incorrect choice). If p-value and b-value have similar prediction performance, the absolute residuals from both predictors should be similar when predicting choice correctness for the same trial. The average difference between the absolute residual of p-value prediction and the absolute residual of b-value over

all rats' trials is approximately 0.0012 (note that the average difference could be maximally 1.0), suggesting that p-value and b-value have similar performance in predicting the next trial's outcome.

### **3.2 Bayesian learning criterion is not always consistent with 6-in-a-row criterion**

A *Bayesian learning criterion* can be used to determine whether or not a rat has learned the reward-associated stimulus in each learning stage. We know all the 93 archival rats' data were collected based on the 6-in-a-row criterion, but would the performance of these rats satisfy a Bayesian learning criterion? We computed the posterior probability of the reward hypothesis corresponding to the reward-associated response pattern on the final trial (when 6-in-a-row correct responses had been made) of each stage for each rat (Figure 3.3). As a reasonable Bayesian criterion, we chose a posterior probability for the reward hypothesis associated with the reward-associated response pattern of 0.95. From these histograms (Figure 3.3), we can see that most rats satisfy this Bayesian learning criterion in the CD stage, but about half or fewer rats did not satisfy the Bayesian criterion in each of the other stages, respectively. This result might indicate that the pattern of the rats' responding is consistent with multiple hypotheses, causing low Bayesian estimate of the correct (*rewarded*) hypothesis. Another possible reason is that the patterns of the rats' initial responding are consistent with non-rewarded response patterns, forcing the posterior probability of the hypothesis associated with correct responding to be very low just before the last six trials.

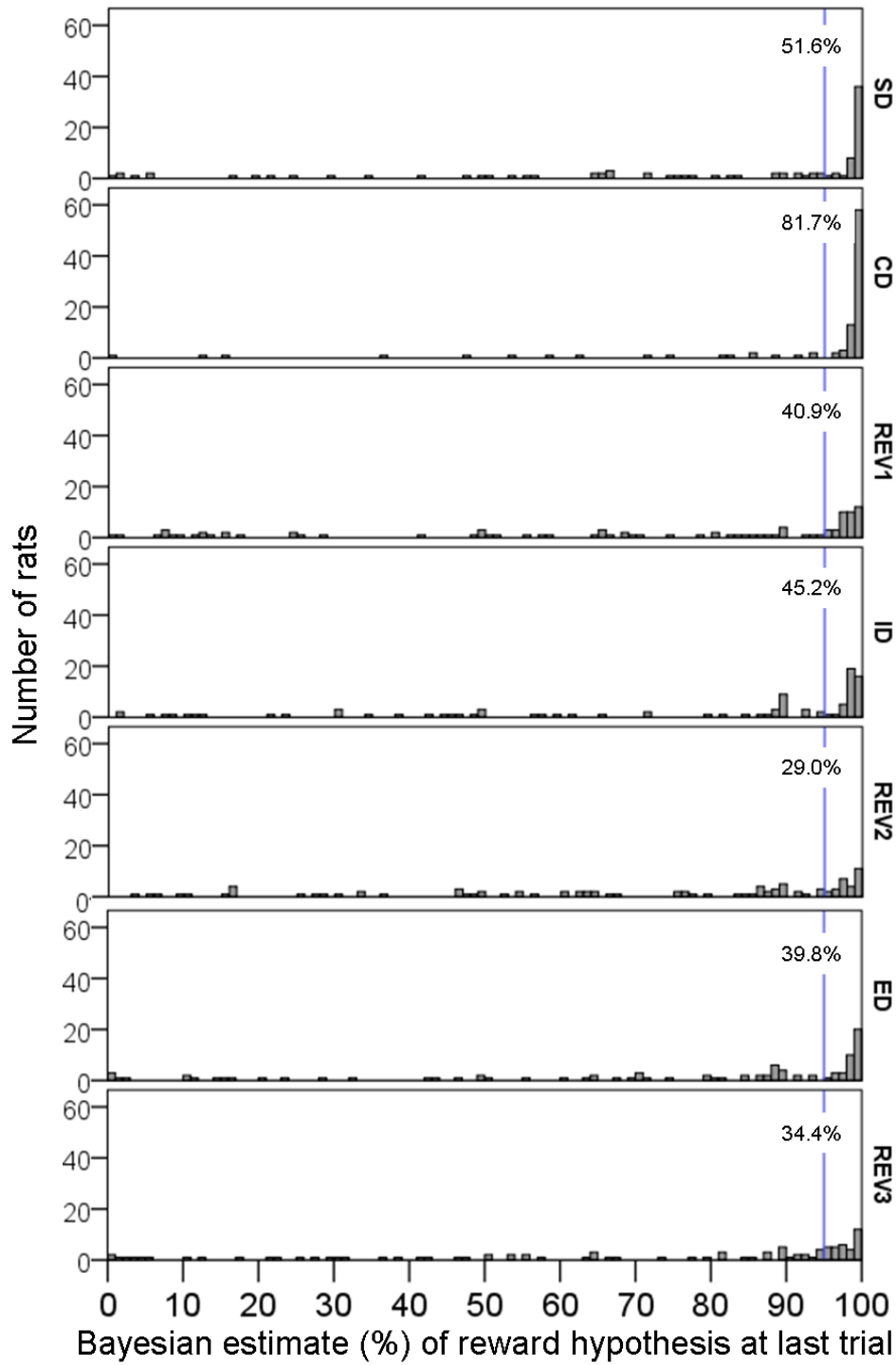


Figure 3.3: histogram of the last-trial Bayesian posterior probability of the hypothesis corresponding to the reward-associated response pattern (called ‘reward hypothesis’ henceforth) for each learning stage. Vertical blue line: the position of the Bayesian threshold 0.95 on the x-axis. By definition, all animals reached the 6-in-a-row criterion on the trial,  $p < 0.016$  via null hypothesis testing. Although the p-values are identical, the varying Bayesian posterior probabilities raise the possibility that learning was not uniform across rats at each stage. Indeed, note that some of the Bayesian posterior probabilities are close to 0. The percentage of rats reaching the Bayesian learning criterion was also shown on the top right for each stage.

### 3.3 Bayesian analysis of individual rat data

Because the Bayesian approach can estimate the probability of each hypothesis for *each* trial in a learning stage, we can analyse the learning process of individual rats for each learning stage. Figure 3.4 (top) demonstrates individual analyses of two rats for the ED stage. Each curve represents the Bayesian probability of a given hypothesis over the course of the trials, except that the black curve represents sum probabilities of two hypotheses from the reward-irrelevant perceptual dimension. From the top left figure (rat ‘06/188’), we can infer that the rat first focused on the reward-irrelevant perceptual dimension (black curve) and then shifted to the other perceptual dimension, using the reward-associated stimulus (the solid blue curve) over the unrewarded stimulus from that dimension (the dashed blue curve). In comparison, the other rat on the right (‘06/190’) probably tried the win-stay spatial response pattern (solid green) to choose bowls from the beginning and lasted for quite a few trials before finally focusing on the reward-associated response pattern. Note that it is difficult to obtain such detailed *individual* analysis using the traditional frequentist approach, by which an estimate of multiple response patterns at each trial is impossible and in general only the statistics of *groups* of data can be obtained.

While we can analyse rats’ individual performances as above, such analysis is descriptive rather than quantitative. To quantitatively represent the individual performance of rats, we proposed to use the correlation between the posterior probabilities of relevant dimension and irrelevant dimension over trials in the ED stage, where the posterior probability of relevant dimension is the sum of the posterior probabilities of the two perceptual hypotheses in the relevant perceptual dimension, and similarly for the posterior probability of irrelevant dimension. Figure 3.4 (bottom) shows the histogram of the correlations between posterior probabilities of relevant and irrelevant dimensions across trials for both lesion and control groups. As expected, the correlations from the above two rats’ data in the ED stage are quite different: one having very strong negative correlation, and the other having a weak and positive correlation (see the two blue arrows in Figure 3.4). Interestingly, the histogram shows

that the distributions of correlations are clearly different between the two groups,  $D_{47,46} = 0.44$ ,  $p < 0.01$  with two sample Kolmogorov-Smirnov test. Most control rats have negative and strong correlations (Figure 3.4 bottom, histogram in light gray); in comparison, most mPFC-lesioned rats have weak correlations (Figure 3.4 bottom, shaded histogram). This may indicate that control rats are more likely focused on perceptual response patterns in the ED stage, which may have led to strong negative correlations between the two perceptual dimension's probabilities. In comparison, lesioned rats may have tried various spatial response patterns more often, such that the posterior probabilities of relevant and irrelevant dimensions would become smaller and more likely have similar trends or no obvious trends across trials, which would then cause either positive correlations or null correlations.



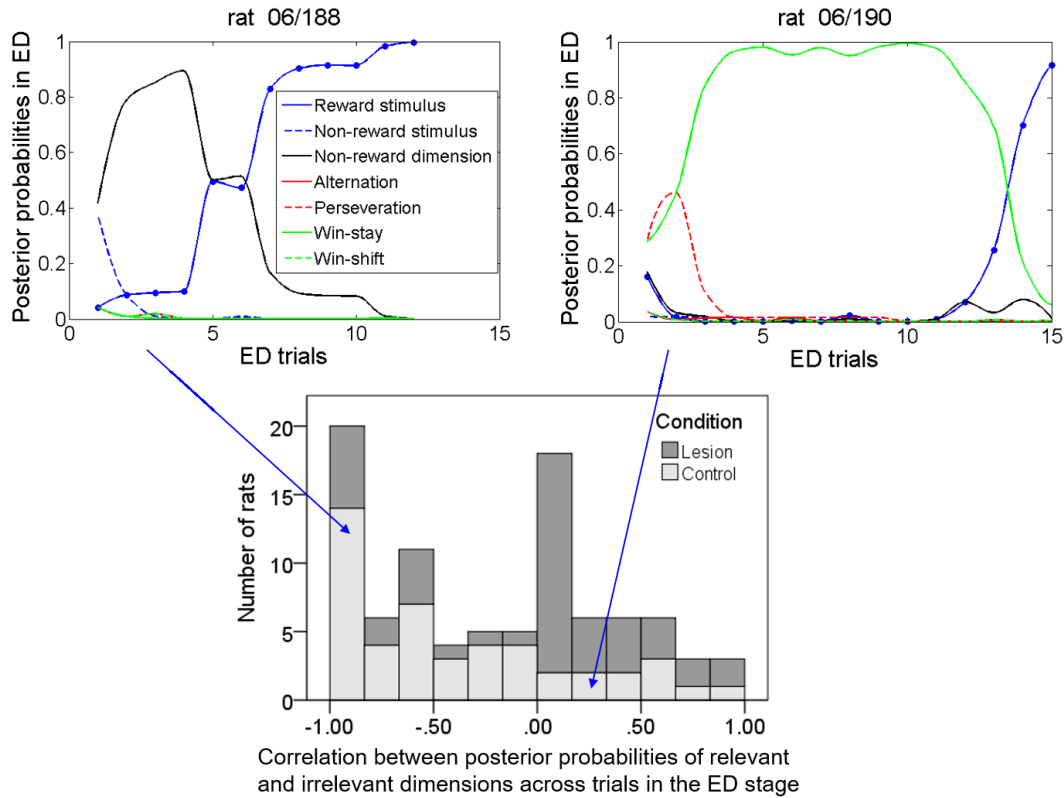


Figure 3.4: Bayesian analysis of individual performance at the ED shift stage, in which normal rats are expected to respond initially to the old perceptual dimension before learning that the previously unattended perceptual dimension is now relevant. One normal rat 06/188 (top left) responded in a way that was consistent with attending to the irrelevant perceptual dimension (black curve) initially and then shifting to the rewarded response pattern (solid blue curve), thus causing a strong negative correlation between Bayesian probability of reward and the irrelevant dimension across trials (see the blue arrow from top left to bottom). In contrast, another normal rat 06/190 (top right) used an apparent win-stay response pattern (solid green) rather than a response pattern according to the previously attended perceptual dimension. Thus, the hypothesis linked with the previously attended perceptual dimension had a posterior probability close to 0, and consequently, a low correlation between Bayesian estimates of the relevant and irrelevant dimensions (see the blue arrow from top right to bottom). Bottom: stacked histogram; around half of the control rats have a strong negative correlation between Bayesian probabilities of relevant and irrelevant dimensions in the ED stage, while the majority (65%) of lesioned rats have weak correlations. Such difference in the correlation between control and lesioned rats suggests that lesioned rats used other response patterns than simply tracking the previously attended perceptual dimension, and thus, the impairment in those rats might not have been due to being unable to shift attentional set per se.

### 3.4 Both lesioned and normal rats tried multiple spatial response patterns

Based on individual analysis, we extracted more information about the rats' response patterns that could not be obtained from the original frequentist approach applied to

these data. We know that rats have strong spatial memory and therefore can use spatial patterns in learning. If a rat used a reward-irrelevant (including spatial) response pattern systematically, the rat would receive reward 50% of the time. At this low rate of reinforcement, it is less likely that the posterior probability of the reward-irrelevant hypothesis will increase as high as 0.95. Therefore, we used a moderately high value (i.e., 0.6 here) to infer if a rat used spatial (therefore reward-irrelevant) response patterns, i.e., if the Bayesian posterior probability of a spatial response pattern is larger than 0.6 in a trial, we would consider that the rat used the spatial response pattern to make choice in the trial. Note that while Bayesian analysis result was used to determine whether a rat tried spatial patterns, the criterion to determine whether a rat tried spatial patterns is non-Bayesian. This is because Bayesian analysis can only provide the probability of trying a spatial hypothesis rather than make a decision whether a rat tried the specific hypothesis. In other words, the decision making here (i.e., whether a rat tried a specific spatial response pattern) is non-Bayesian. With this criterion (i.e., a rat tried the spatial response pattern if the posterior of the spatial hypothesis is larger than 0.6), we found that each rat may have tried two to four spatial response patterns. Table 3.1 shows that each spatial response pattern appears to have been used by the majority of both lesioned and control rats. In particular, spatial perseveration was used by most lesioned and control rats. However, Chi-Square Goodness-of-Fit Test shows there is no significant difference in the proportion of lesioned rats using different spatial response patterns (Chi-Square(3)=1.08, n.s.; Table 3.1, first row), and similarly for the control rats (Chi-Square(3)=3.60, n.s.; Table 3.1, second row). On the other hand, the frequencies of rats using spatial alternation are significantly different between groups (Chi-Square(1) = 9.40,  $p < 0.01$ ; Table 3.1, second column), while the frequencies are not significantly different between groups for each of the other three spatial response patterns.

Although the extensive use of spatial response patterns complicates the analysis of the attentional set-shifting task trying to assess rats' performance on ID and ED set-shifting ability, it may also provide additional information about rats' learning. For

instance, the use of spatial response patterns might be a very good indicator that the rat doesn't have confidence about where the reward can be found.

Table 3.1: number (%) of rats using each spatial response patterns in the 7-stage task. We consider that a rat used a spatial response pattern if the posterior probability of the corresponding spatial hypothesis was larger than 0.6 on any learning trial of the task.

	Alternation	Perseveration	Win-stay	Win-shift
Lesioned rats (out of 46)	<b>41</b> <b>(89%)</b>	44 (96%)	35 (76%)	39 (85%)
Control rats (out of 47)	<b>29</b> <b>(62%)</b>	42 (89%)	30 (64%)	39 (83%)

### 3.5 Lesioned rats tried more spatial trials in ED stage than normal rats

We have demonstrated that rats probably used spatial response patterns to choose bowls during learning. Here we also would like to know how *long* they tried spatial response patterns and whether lesion is associated with the length of spatial patterns, particularly in the ED stage. From previous studies, we already know mPFC-lesioned rats take longer to learn the ED stage, as seen from the first pair of bars in Figure 3.5. Figure 3.5 (middle pair of bars) also shows that lesioned rats use spatial response patterns on significantly more trials;  $m=1.62$ ,  $sd=2.36$  for control group,  $m=6.41$ ,  $sd=5.59$  for lesion group,  $t(91)=5.41$ ,  $p<0.001$ ,  $d=0.57$ . Interestingly, when we exclude such spatial trials (where rats chose bowls using spatial response patterns) from the total trials of the ED stage for each rat, we found there is no significant difference between both groups of rats on trying other patterns;  $m=11.34$ ,  $sd=4.47$  for control group,  $m=13.45$ ,  $sd=6.05$  for lesion group,  $t(91)=1.92$ ,  $p=0.058$ , ns. This suggests that the difference in ED trials between the two groups is mainly from the difference in the number of spatial trials used by both groups.

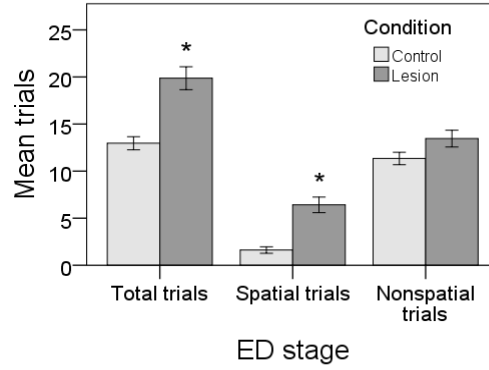


Figure 3.5: number of spatial trials in ED is different between two groups.

### 3.6 Lesioned rats tried more spatial trials in reversal stages than normal rats

Besides the ED stage, we also compared the length of spatial response patterns for other stages in the 7-stage task. Two-way ANOVA shows that the number of spatial trials used by lesioned rats are significantly higher than that in control rats (Figure 3.6),  $F(1,637)=50.51$ ,  $p<0.05$ , partial  $\eta^2=0.073$ . Similarly, the proportion of spatial trials (relative to the total number of trials within each stage) used by lesioned rats are also significantly higher than that in control rats (Figure 3.7),  $F(1,637)=35.99$ ,  $p<0.05$ , partial  $\eta^2=0.053$ .

Independent two-sample  $t$ -tests with Dunn-Sidak correction showed that the number of spatial trials in each reversal stage is significantly higher in lesioned rats than that in control rats (Figure 3.6);  $t(91)=-2.20$ ,  $p<0.05$ ,  $d=0.23$  for REV1,  $t(91)=-2.64$ ,  $p<0.05$ ,  $d=0.28$  for REV2,  $t(91)=-3.17$ ,  $p<0.05$ ,  $d=0.33$  for REV3. Similarly, an independent two-sample  $T$ -test with Dunn-Sidak correction shows that the proportion of spatial trials from both REV2 and REV3 are significantly higher in lesioned rats than those in control rats (Figure 3.7),  $t(91)=-2.31$ ,  $p<0.05$ ,  $d=0.24$  for REV2 and  $t(91)=-3.01$ ,  $p<0.05$ ,  $d=0.32$  for REV3, while the proportion of spatial trials from REV1 is not significantly different between the two groups,  $t(91)=-1.38$ ,  $p=0.23$ , ns. These results suggest that lesioned rats used longer spatial response patterns in learning than control rats, particularly in the latter two reversal stages.

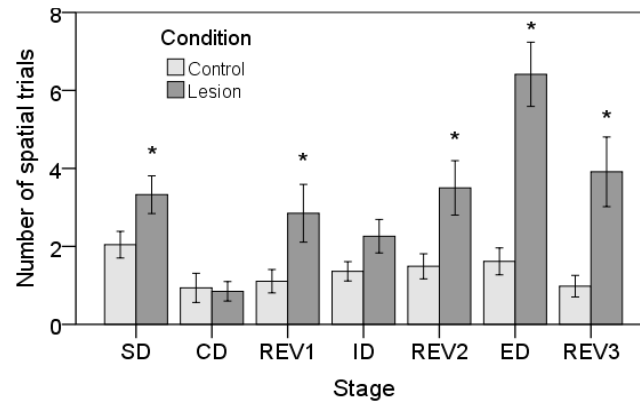


Figure 3.6: number of spatial trials within each stage for both control and lesion groups. Lesioned rats used significantly more spatial trials than control rats in all three reversal stages as well as the ED stage.

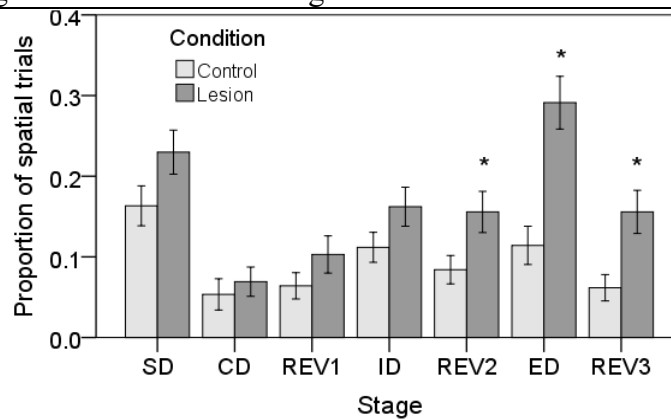


Figure 3.7: the proportion of spatial trials relative to the total number of trials within each stage for both control and lesion groups. Lesioned rats used a significantly higher proportion of spatial trials than control rats in the latter two reversal stages (REV2 and REV3), as well as the ED stage.

### 3.7 Association between length of spatial pattern and learning

Considering rats have widely used spatial response patterns to choose bowls in learning, it would be interesting to investigate the association between the number of spatial trials and the length of learning for each stage. Table 3.2 (first row) shows that there are strong correlations between the number of spatial trials and the total number of trials within each stage for the lesion group. That means, the more trials where lesioned rats used spatial response patterns to choose bowls within a stage, the more trials spent to finish that stage. In comparison, the correlations are weaker in the control group than those in the lesion group (Table 3.2, second row). Also see Figure 3.8 for examples about the correlations for each group.

Table 3.2: correlations between the number of spatial trials and the total number of trials within each stage or with the combination of all stages, for both the lesion group (46 rats) and control group (47 rats). All the correlations are significant ( $p<0.05$ ). Besides Pearson's correlation coefficient  $R$ , Spearman's  $\rho$  (in parentheses) was also reported due to heteroscedasticity of spatial trials in each stage.

	All stages	SD	CD	REV1	ID	REV2	ED	REV3
Lesion group	0.780 (0.649)	0.805 (0.777)	0.751 (0.614)	0.807 (0.655)	0.794 (0.619)	0.761 (0.620)	0.680 (0.664)	0.877 (0.702)
Control group	0.464 (0.441)	0.680 (0.622)	0.616 (0.520)	0.508 (0.293)	0.610 (0.568)	0.530 (0.551)	0.367 (0.366)	0.425 (0.481)

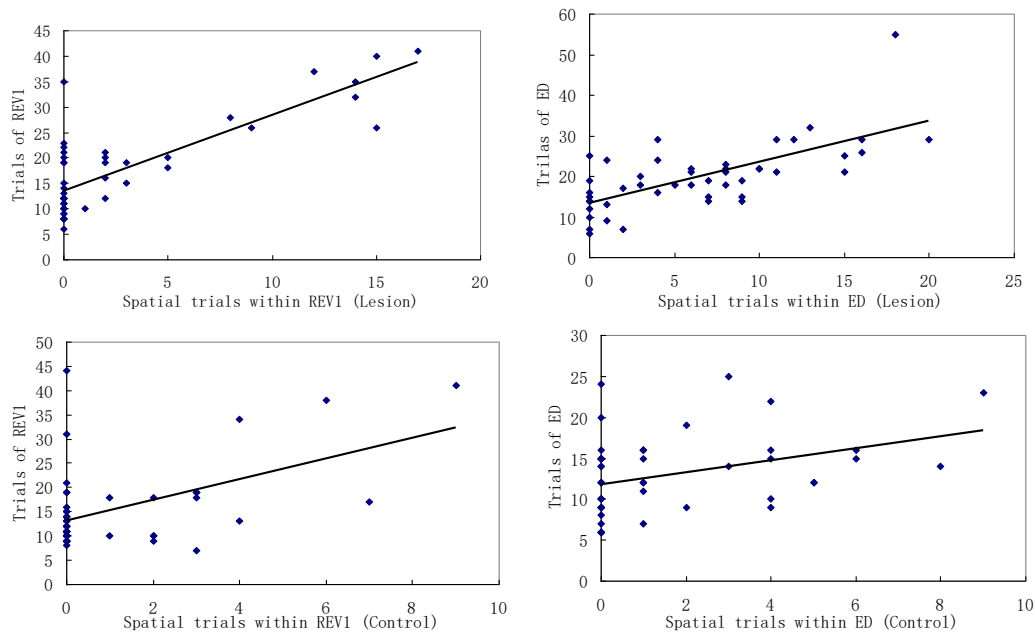


Figure 3.8: scatter plots to show the correlations between the number of spatial trials and the total number of trials to finish the REV1 stage (first column) and ED stage (second column), for 46 lesion rats (first row) and 47 control rats (second row) respectively. The correlations are weaker in the control group than in the lesion group. See Table 3.2 for the complete set of correlations.

### 3.8 First spatial response patterns at the ED are different between two rat groups

With a Bayesian analysis, not only the length of spatial response patterns, but also the types of spatial response patterns can be analysed. Specifically, the type of first spatial hypotheses that rats use in the ED stage can be assessed. Figure 3.9 shows that the frequency distribution of first spatial response patterns is significantly different between control and lesion groups, Chi-square(4)=16.53,  $p<0.01$ . Lesioned rats are inclined to use spatial perseveration before other spatial patterns (including spatial alternation, win-stay, and win-shift) in the ED stage, while the majority of control rats

did not try any spatial patterns in the ED stage.

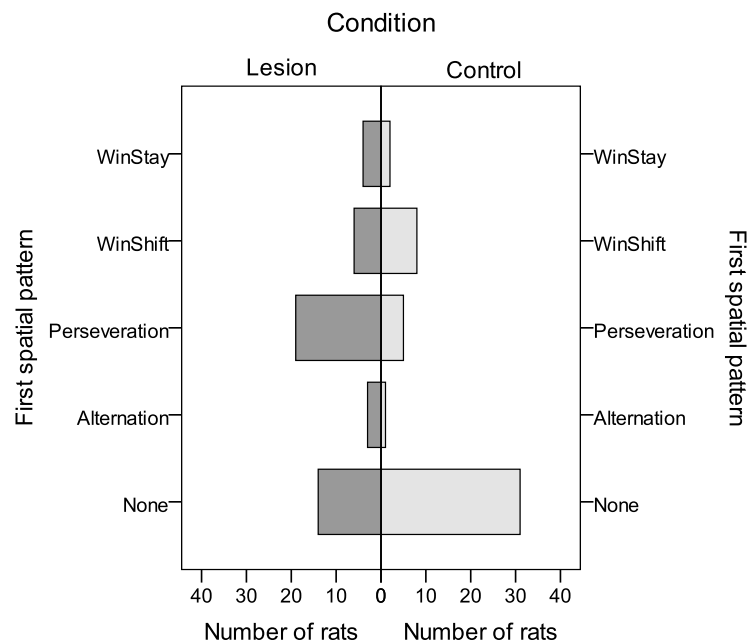


Figure 3.9: frequency of first spatial patterns used in both lesion and control groups. Most control rats did not use any spatial response patterns, while lesioned rats predominantly used spatial perseveration in choosing bowls. The total number of rats are similar between two groups, i.e., 47 control rats and 46 lesioned rats.

### 3.9 Overtraining reversal effect

When an animal continues to be trained (e.g., overtrained over 50 trials) after reaching a learning criterion (e.g., 6-in-a-row) in one learning stage, the animal would more likely learn faster in the subsequent reversal learning stage compared to those animals without being over-trained. This phenomenon is called the overtraining reversal effect (Reid, 1953; Mackintosh, 1974). The overtraining reversal effect is associated with ‘difficult’ learning, i.e., learning where there are distractions (irrelevant dimensions). So what overtraining does is to strengthen attentional set, by providing more experience of relevant vs irrelevant dimensions. During a reversal where there is no overtraining, a subject would likely have just a weak (or no) attentional set, and the stimuli from the irrelevant dimension would be partially reinforced before reversal (Mackintosh, 1974). In this case, in the reversal stage, once the subject found the previously correct stimulus becomes incorrect, the partial reinforcement of the irrelevant-dimension stimuli would make the subject likely choose to try an

irrelevant-dimension stimulus over the previously incorrect stimulus within the relevant dimension. However, if the subject is over-trained before the reversal, the subject would likely attend to the relevant dimension, knowing that the stimuli from the irrelevant dimension are irrelevant to reward. In other words, the partial reinforcement of irrelevant-dimension stimuli would be largely reduced by the overtraining. In this case, in the reversal stage, once the subject found the previously correct stimulus does not correspond to reward, the subject would more likely try the previously incorrect but currently correct stimulus within the relevant dimension rather than irrelevant-dimension stimuli. As a result, overtraining would make the subject take fewer errors to learn the reversal.

While the rats performing the 7-stage task were not explicitly over-trained when judged by 6-in-a-row criterion, some rats may have received more training than necessary if judged by the Bayesian learning criterion. Here, we explored whether an apparent overtraining reversal effect exists using the new Bayesian learning criterion on the archival rats' data.

Although rats were not explicitly overtrained, the rats received more training on the correct stimulus prior to the first reversal stage REV1 than prior to the latter two reversal stages REV2 and REV3, because food reward was associated with the same stimulus in both the SD and CD stages before REV1. Here, the degree of training before the REV1 stage is represented by the number of trials of which the Bayesian probability of the reward-associated hypothesis is larger than 0.95 in the previous SD and CD stages. Similarly, the degree of training before REV2 and REV3 was collected from their preceding stages, ID and ED, respectively. Figure 3.10 (first row) shows there is a slight trend of overtraining reversal effect for REV1 stage in both lesioned and control groups. In comparison, such a slight trend decreased or disappeared for the other two reversal stages particularly for the control group (Figure 3.10, second and third rows), probably because there is relatively less training before those two reversal stages.



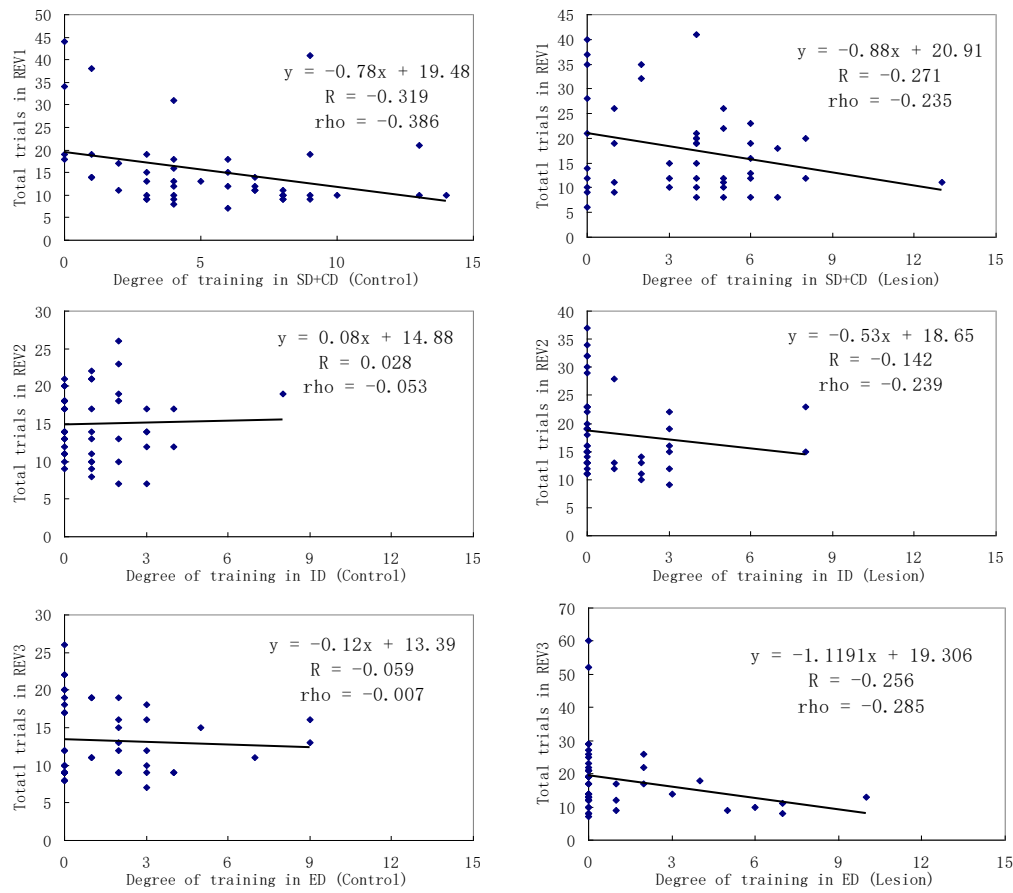


Figure 3.10: correlations between the trials of a reversal stage (y-axis) and the degree of training in the previous stage (x-axis), for both the control group (left column, 47 rats) and lesion group (right column, 46 rats). There is slight overtraining reversal effect in the first reversal stage REV1 (first row) for both groups. In each figure, R is Pearson's correlation coefficient, and rho is Spearman's rank correlation coefficient.

## **4 Discussion on Bayesian analysis of rat data**

Building on the initial exploration of a Bayesian analysis in an attentional set-shifting task in my undergraduate project, I have extended and refined the Bayesian approach in both the likelihood function and the prior function. Compared to the traditional frequentist approach, Bayesian analysis provides us with detailed information about what response patterns may have been tried in each learning trial for each rat. Based on the individual analysis, we found that all rats used intermittent spatial patterns of responding while they were learning the stimulus-reward associations, and mPFC-lesioned rats used more spatial trials than control rats in ED and reversal stages. In addition, the appropriateness of the Bayesian learning criterion was validated on real 7-stage data, although the Bayesian learning criterion does not always match with the frequentist 6-in-a-row criterion in determining the time point at which rats learned the correct stimulus-reward association.

### **4.1 Bayesian approach**

The observations regarding whether the rat chooses to dig after encountering only one, or both, bowls has been appropriately embedded into the likelihood function. If a rat is using a spatial response pattern to choose bowls, it would know which bowl to dig in without observing stimuli in the bowls, and therefore, it would more likely directly go to the to-be-dug bowl. As a result, the rat would more likely dig in the first bowl it encountered. In comparison, if a rat is using a particular perceptual response pattern to choose bowls, the rat cannot decide to choose which bowl to dig in before observing the stimuli in bowls. This is because the stimulus associated with the particular perceptual response pattern could be in either bowl. As a result, if the rat rejects the first bowl it encounters, it is more likely using a perceptual response pattern to choose bowls. For the prior function, a lower-bound prior value is pre-defined for all hypotheses to match the intuition of having a minimum confidence for any hypothesis. With these refinements of the Bayesian analysis from my undergraduate thesis, it is expected that the posterior probability of each hypothesis in each trial is more

accurate than the Bayesian analysis without using the one/both bowls observation. This has been partially confirmed with simulated rat data (see Section 5.4).

Recently, Lloyd and Leslie (2013) proposed a complex Bayesian model which explicitly represented rats' decision-making processes (e.g., decide which response pattern to use for bowl choice in each trial) in order to *simulate* rat behaviour in reversal learning. They modelled the rats' learning and decision-making processes from the *perspective of the animals*. Therefore, their model could help people understand the neural mechanism of how rats learn and make decisions in reversal learning. Similarly, Rygula et al. (2015) applied an existing reinforcement learning model to estimate the probability of marmosets choosing one of two stimuli in simple discriminative learning and reversal learning. Costa et al. (2015) developed a complex Bayesian model to explicitly estimate the probability that rhesus monkeys reversed choice strategy based on the monkeys' choices at each trial in a reversal learning task. Both Rygula's and Costa's methods are from the perspective of the animals and can only estimate the probability of two task-specific hypotheses.

In contrast, our Bayesian approach is used to *describe* rats' patterns of responding from the *perspective of experimenters*, not only in reversal learning but also for both ID and ED shift stages. In essence, the approach presented here can be considered a kind of template matching in which the Bayesian posterior probability estimate quantifies the degree to which the empirical response pattern (i.e., the observation) matches the idealised pattern associated with a given hypothesis. Our approach is based on the pure Bayes rule and was used to analyse the archival data of rats' behaviour and to estimate the probability of multiple ( $>2$ ) spatial and perceptual hypotheses in each learning trial for each rat, not only for simple discrimination but also for compound discrimination learning. By contrast, the other methods do not attempt to analyse data and learning behaviour, but were mainly developed to simulate rats' reversal learning behaviour, or estimate animals' internal decision making. Overall, our Bayesian approach was developed for a unique and different

research aim.

## 4.2 Bayesian learning criterion

The most direct application of Bayesian analysis is to create a new learning criterion based on the estimated Bayesian probability (i.e., posterior probability) of the reward-associated perceptual hypothesis for each discriminative learning stage. When a rat has learned the correct association between the stimuli in the bowls and reward, the Bayesian probability of the reward-associated perceptual hypothesis in general can quickly increase to a high value (e.g., 0.95) in a few trials. Even in the *false negative* conditions under which the rat can be judged to have learned the discrimination but the 6-correct-choice-in-a-row criterion is not satisfied (e.g., 10-out-of-12-correct-choice, Brooks et al., 2012), the Bayesian probability will tend to reach a high value (see section 5.3 for empirical evidence). The moderate correlation between the p-value (used for 6-in-a-row criterion) and the b-value (used for Bayesian criterion) suggests that b-value broadly agrees with p-value but is partially independent of it, raising the possibility that the Bayesian criterion provides different, and perhaps more accurate, information about the subjects' response patterns (Section 3.1). Therefore, by comparing the Bayesian probability of the reward-associated hypothesis with the pre-set high threshold value (0.95) in each learning trial, we can determine whether or not the rat has learned the stimulus-reward association in a timely manner.

Theoretically, the Bayesian learning criterion is also a better choice when there are more than two hypotheses to be tested simultaneously. It is not optimal to accept the hypothesis that rats use the reward-associated stimulus to choose bowls by just rejecting the null hypothesis that rats use random-guess to choose bowls, because there are other reward-irrelevant spatial and perceptual response patterns that are plausible. This means the traditional null-versus-alternative hypothesis testing is severely limited as a means to decide the point at which a rat has learned the stimulus-reward association. Instead, the probability of every plausible reward-irrelevant hypothesis needs to be estimated and rejected in order to accept the reward-relevant

hypothesis. This can be achieved by using a Bayesian approach as presented here.

With the archival rats' 7-stage task data, we found that about half or fewer rats did not satisfy the Bayesian learning criterion in every stage, except the CD stage. In the CD stage, the reward-associated stimulus is not changed compared to the previous SD stage. Therefore, rats can often easily focus on the same reward-associated stimulus from the beginning of the CD stage. Consistently, in most rats the Bayesian criterion would have been satisfied before fulfilling the 6-in-a-row criterion within the CD stage. In this case, the 6-correct-trials-in-a-row criterion (Birrell & Brown, 2000; Chase et al., 2012) is stricter than the Bayesian criterion, and this would suggest the possibility of rat training that is more than necessary. In contrast, in the reversal learning stages, rats need to first suppress the tendency to choose bowls using the previously rewarded (but currently un-rewarded) stimulus and then change to establish a new association between a previously un-rewarded stimulus and the food reward. Because of this difficulty, rats often take more trials, and try various incorrect hypotheses before finding the new correct response pattern. To make the new association stable enough, rats may need relatively more trials to strengthen the new association, and the Bayesian analysis presented here suggests there might have been under-training in reversal stages. In addition, for the rats' data which did not satisfy Bayesian criterion in the ID and ED stages, the rats may happen to respond correctly (i.e., a false positive) in last six trials while actually using alternative response patterns for bowl choice. In this case, the Bayesian criterion suggests more trials would be necessary to rule out such alternative and incorrect hypotheses. The fact that different criteria – Bayesian and 6-correct-trials-in-a-row – differ in either direction with respect to whether learning has occurred indicates that this is not simply a difference in the 'strictness' of the criterion but rather about what can be inferred about cognition from analysis of patterns of behavioural responses.

### **4.3 Bayesian analysis of individual rat's performance**

Besides potentially using Bayesian methods to define a learning criterion, *individual*

analysis of rats' task performance is another advantage of the Bayesian approach. More specifically, the Bayesian approach can determine which response patterns have been tried to choose bowls in each stage for each rat. To the best of our knowledge, this is the first time a detailed learning profile has been estimated for each rat run in an attentional set-shifting task. Such detailed individual analysis could also be obtained by testing multiple null hypotheses with the frequentist approach. However, the frequentist approach is vulnerable to false positives (and false negatives) in evaluating multiple null hypotheses over multiple trials. As a result, the frequentist approach has not been used to obtain detailed individual analysis of rat's performance, but only to obtain the statistics of *groups* of rats' data (Birrell & Brown, 2000; Roberts et al., 1988). Additionally, individual analysis provided by the Bayesian approach might be useful when the sample size is low, e.g., because of the financial and ethical costs of using nonhuman primates (e.g., Roberts et al., 1988), or in case studies of neurological or psychiatric patients (e.g., Prentice et al., 2008; Lange et al., 2016).

Furthermore, we proposed a quantitative measure, i.e., the correlation between the posterior probabilities of relevant dimension and irrelevant dimension over trials within a stage, to quantify the transition from responding to the old dimension to the new one. When a rat focuses on perceptual response patterns for bowl choice, the correlation would be negative and strong particularly at the start of the stage. Instead, if a rat tries various spatial response patterns during learning, the correlation would be more likely positive and/or weak. We have found that the distributions of such correlations over rats are different between intact and neurologically impaired rats. Most control rats have strong and negative correlations in the ED stage, while most mPFC-lesioned rats have weak or positive correlations in ED. This suggests that most control rats may have focused on perceptual dimensions during ED learning, while most mPFC-lesioned rats may not.

This quantitative measure could be potentially used to explore the degree of set-formation for individual rats. Since the purpose of the ED stage is to explore the rats'

ability to shift from the previously formed attentional set to a new attentional set, the ED stage can only test set-shifting ability if a set has been well-formed. To date, the significant difference in the number of learning trials between the ID and the ED stage at the *group* level is used to support the idea that rats have formed attentional set before the ED stage (Birrell & Brown, 2000; Brown & Tait, 2016; Roberts et al., 1988). The rationale behind this traditional criterion is that old learning would interfere with new learning as the context changes such that rats would require more effort (i.e., an increased number of trials) to finish the new learning if attentional set is formed in the old learning. However, the group-level difference between ID and ED trials does not mean *each* rat takes more trials in the ED stage compared to in the ID stage. Actually, six out of 46 lesioned rats and 17 out of 47 control rats took either equivalent or fewer trials in the ED stage compared to the ID stage. More importantly, the fact that a rat has more ED trials than ID trials does not necessarily mean that the rat had formed the attentional set. If a rat has a well-formed attentional set before the ED stage, the rat should choose bowls based on the stimuli within the previously relevant but currently irrelevant dimension at the beginning of the ED stage, increasing the probability of the irrelevant dimension and decreasing the probability of the relevant dimension that is associated with reward. Also, in the last six trials within the ED stage where rats almost likely tried the rewarded stimulus response, the probability of the relevant dimension often increases, while the probability of irrelevant dimension response will be kept at small values or even further decrease. Considering the contrary changes of two dimensions' probabilities at the beginning and the end of the ED stage, any rat with good set-formation would probably have a negative and often strong correlation between the two dimensions' probabilities within the ED stage. In contrast, rats without robust set-formation would likely lead to a weak or positive correlation within the ED stage (see Figure 3.4 as an example). This quantitative measure could potentially be used to determine whether each rat has a well-formed attentional set for further data analysis about ED performance.

#### **4.4 Rats tried spatial response patterns in both control and lesioned groups**

Based on the individual analysis of each rat's learning process, we found that both lesioned and normal rats tried multiple types of spatial response patterns in the 7-stage task, with spatial perseveration slightly more common than the other three spatial response patterns. Foraging for food is an inherently spatial task and rats are very adept in mapping their environment (Geva-Sagiv et al., 2015). In the natural environment, spatial memory might allow the rat to forage more efficiently by avoiding reliance on perceptual cues that are proximal to the potential food sources. In effect, if the rat has a spatial map, it might not need to know the perceptual characteristics of a given location. Furthermore, in this task, the rat must dig in one of two bowls on each trial, even before it knows which one contains food. The initial 'look and see' that characterises at least the first trial, and likely a few more, may be based on a spatial choice, but that does not mean that the rat is (necessarily) not attending to the perceptual aspects of the bowl in which it is digging. Obviously, it is never possible to 'know' what the rat knows or thinks. We can only observe behaviour or neural information processing. However, the rapidity with which a rat changes its behaviour from following a spatial behavioural pattern to responding consistently to the correct stimulus, strongly suggests that it is learning (acquiring information) about the correct stimulus even while making initial bowl selections based on spatial properties.

#### **4.5 Lesioned rats tried spatial patterns for more trials than control rats**

Again, based on individual analysis of each rat's performance, we also found that mPFC-lesioned rats spent more trials using spatial response patterns for bowl choice than control rats, and interestingly, the number of remaining ED trials, after excluding spatial trials, was similar for the lesion and control groups. In the ED stage, rats need to disengage from the old attentional set (i.e., the previously relevant but currently irrelevant dimension) and then learn a new attentional set. During the disengagement process, rats would probably focus on the two stimuli of the previously rewarded perceptual dimension. After disengagement, rats may try various response patterns,



including spatial response patterns, in order to learn the association between food reward and the new dimension. Therefore, the result we found here may indicate that the lesion of mPFC may mainly affect rats' ability to attend to or associate with the new reward-relevant dimension (e.g., mPFC-lesioned rats might react to uncertainty differently by requiring more evidence before they commit to reward hypothesis), rather than affect disengagement from the old dimensional set.

This is inconsistent with the previous study's result showing that mPFC-inactivation mainly caused perseveration errors in the shift stage (Floresco et al., 2008). Such inconsistency may come from multiple factors. The previous study is based on a lever-pressing task which requires rats to shift from visual cue discrimination to position discrimination, whereas our data is from the 7-stage task which requires rats to shift from a discriminating medium to discriminating odours (or from odours to medium). In the lever-pressing task, the same visual stimuli were still there and spatial response is relevant with reward in the shift stage, while the medium and odour stimuli in the ED stage of the 7-stage task are novel and have not appeared in previous stages, and spatial response patterns are never usable to solve the discrimination in the 7-stage task. In the lever-pressing task, it is expected that rats would always get perseverative errors in the shift stage because the previously reward-relevant stimulus (not just dimension) was still present in the shift stage. The differences in types of discrimination learning and in novelty of stimuli in the ED stage between the two tasks may somehow interact with the effect of mPFC lesion on ED shift, causing different difficulties during the rat's learning. Another factor may come from the data collection procedure. In the first four trials of each learning stage in the 7-stage task, each rat was allowed to dig the other bowl if it dug in the first, unrewarded bowl (Tait et al., 2014). This special step may make rats more quickly disengage from the previously relevant but currently irrelevant dimension, thus suppressing the perseverative errors that rats would have made. From this perspective, the lack of mPFC lesion effect on perseverative errors from the Bayesian analysis may be potentially from the special data collection step during the first four trials in

the ED stage. Further study without this step in data collection with the 7-stage task may help clarify whether mPFC lesion really affects perseverative errors or not.

Interestingly, besides the ED stage, we also found that lesioned rats spent more trials using spatial response patterns than control rats, particularly in the latter two reversal stages in the 7-stage task. Together with the above finding for the ED stage, this finding suggests that lesions of mPFC may affect rats' ability to not only attend to the previously irrelevant perceptual dimension in the ED shift, but also attend to the previous unrewarded but currently rewarded stimulus of the same perceptual dimension in reversal learning. This again is not consistent with previous findings that mPFC lesion only affects rat's or mice's ED shift performance (Birrell & Brown, 2000; Bissonette et al., 2008; Brown & Tait, 2016). However, initial analysis of our large set of rats' data (46 lesioned and 47 controls) did show that lesioned rats spent significantly more trials completing not only the ED stage but also the latter reversal stages than control rats. Therefore, that previous studies did not find the effects of mPFC lesions on reversal learning might have been due to insufficient statistical power given the modest effect size. Alternatively, there might be other noise factors which caused such inconsistency between our findings and previous studies which also used the 7-stage task for data collection, such as the location and size of the lesions. Therefore, further study particularly with a large dataset is necessary to clarify how mPFC lesion can really potentially affect reversal learning in compound discrimination.

Following the investigation of spatial response patterns, we also found that there are strong correlations between the number of spatial trials and the total number of trials within *each* stage for the lesion group. This is reasonable, because when rats took longer to find the correct hypothesis, they would likely use various spatial response patterns in more learning trials. Unexpectedly, the same rationale does not seem valid for control rats, because we found the correlations between the number of spatial trials and stage trials are clearly weaker, but still statistically significant. One possibility is

that normal rats may more flexibly change hypotheses during learning, such that spatial response patterns used by normal rats are more difficult to estimate than those used by lesioned rats.

#### **4.6 First spatial hypothesis in ED**

Bayesian analysis of individual rats' performance can also help answer certain questions of special interest from researchers. One interesting question is about the type of first spatial response pattern rats used in the ED stage. With Bayesian analysis, we found lesioned rats more likely tried spatial perseveration than the other three spatial response patterns, while most control rats did not try any spatial response pattern at all in the ED stage. The limited number of control rats which tried spatial response patterns showed a slight preference for win-shift compared to the other three spatial response patterns in bowl choice. Since spatial perseveration requires less cognitive burden than win-shift, and the low use of spatial response patterns in most control rats may come from quick (flexible) changes between different response patterns for bowl choice, the above finding provides a different type of evidence to support the idea that control rats are more cognitively flexible than lesioned rats.

#### **4.7 Overtraining reversal effect**

Besides the powerful applications of the proposed Bayesian approach in various individual, and subsequent group-level, analyses, we also tried to explore the traditional overtraining reversal effect (Reid, 1953; Mackintosh, 1974) in the 7-stage task data, where the degree of apparent overtraining before a reversal stage is represented by the number of trials after reaching the Bayesian criterion. Overall, we did not find an overtraining reversal effect, although there is a slight trend for the first reversal stage (REV1) in both groups. This finding seems to be inconsistent with the traditional overtraining reversal effect (Reid, 1953; Mackintosh, 1974). However, the traditional overtraining reversal effect is obtained based on explicit overtraining of rats after reaching a learning criterion (e.g., 6-correct-in-a-row). In comparison, for the archival data we analysed, each rat was stopped immediately after meeting the 6-correct-in-a-row criterion, therefore largely limiting the degree of potential

overtraining. As a result, it is plausible not to find the overtraining reversal effect from the less likely overtrained rats' data. This rationale is supported by the slight trend of apparent overtraining reversal effect in the first reversal stage, where rats may have received slightly more apparent overtraining before the reversal stage (than those before the latter reversal stages) because the same reward-associated stimulus was used in both SD and CD stages.

#### **4.8 Summary**

In summary, all the above results and discussions have shown the Bayesian approach can extract richer information from set-shifting task data, particularly based on individual analysis of rats' performance. What is important is that these results are not just about the data from 93 rats undertaking the 7-stage task. The results are about how we infer what approach animals take in learning. We used this large dataset and results to demonstrate the general applications of Bayesian analysis that can be used in other learning situations.

## **5 Bayesian analysis of simulated data**

In the 7-stage task, we never know for certain the response patterns a given rat uses to choose bowls at each trial. We can only estimate the probability of corresponding hypotheses at each trial by the proposed Bayesian approach. Therefore, one fundamental question arises: does a higher Bayesian estimate of a hypothesis really indicate the rat is using the corresponding response pattern to make its choices?

To answer this question, virtual rat behaviour can be simulated and then analysed by the proposed Bayesian approach. The basic idea of simulation is to assume that we know what sequence of response patterns a virtual rat would use to learn a virtual stage (no matter what the stage is). Then together with the real pre-determined bowl designs (which bowls contains which stimuli in each trial), any sequence of known response patterns would generate simulated (virtual) stage data. Once we get simulation data, we can apply the Bayesian analysis to the simulated data to estimate Bayesian probabilities for each hypothesis associated with a specific response pattern at each (virtual) trial in the simulation data. The Bayesian analysis result can then be compared with the ground-truth response patterns underlying the simulated data to evaluate the validity of the Bayesian approach.

### **5.1 Individual analysis on simulated data**

One concern regarding the Bayesian analysis of individual rats' performance is whether a high Bayesian probability of a spatial hypothesis really indicates that the rat used the corresponding spatial response pattern to dig in the bowls. To answer this question, we assume we know virtual rats would use a series of spatial response patterns followed by six reward-relevant hypotheses to learn a stage (Table 5.1).

Table 5.1: a series of known ‘ground-truth’ response patterns used to generate simulation data. ‘Random’ means the virtual rat would randomly dig in either the left or the right bowl.

Trials	1	2-5	6-11	12-15	16-19	20-25
Hypothesis	Random	Spatial perseveration	Win-stay	Spatial alternation	Win-shift	Rewarded stimulus

Based on the ground-truth response pattern series and the same pre-determined bowl designs as used in real data collection, a simulation observation can be generated (Figure 5.1). For example, if *spatial perseveration* is the ground-truth response pattern for the 5<sup>th</sup> trial, then the side choice on this trial will be the same as that in the previous (4<sup>th</sup>) trial, e.g., the left side in Figure 5.1. If the reward is in the chosen bowl, the rat would get the reward; otherwise not (as shown in Figure 5.1). This will generate an observation about which bowl contains which stimuli, which bowl contains a food reward, which bowl was chosen, and whether the rat got a reward or not for the 5<sup>th</sup> simulated trial.

So far, only the observation about 1<sup>st</sup>/2<sup>nd</sup> bowl choice (i.e., digging in the first bowl the rat approached, or rejecting that and investigating the second bowl before making a response) remains to be simulated for each trial. To generate the 1<sup>st</sup>/2<sup>nd</sup> observation for the k<sup>th</sup> trial, if the k<sup>th</sup> response pattern in the known response pattern series is a spatial one, the virtual rat would dig in the 1<sup>st</sup> bowl with probability 0.9 and or reject the 1<sup>st</sup> bowl and investigate the 2<sup>nd</sup> bowl with probability 0.1 (Figure 5.1); otherwise, if the k<sup>th</sup> response pattern is perceptual, the virtual rat would dig in, or reject, the 1<sup>st</sup> bowl equally with probability 0.5. This way to generate 1<sup>st</sup>/2<sup>nd</sup> observation is consistent with the parameter setting in the likelihood function of the Bayesian model.

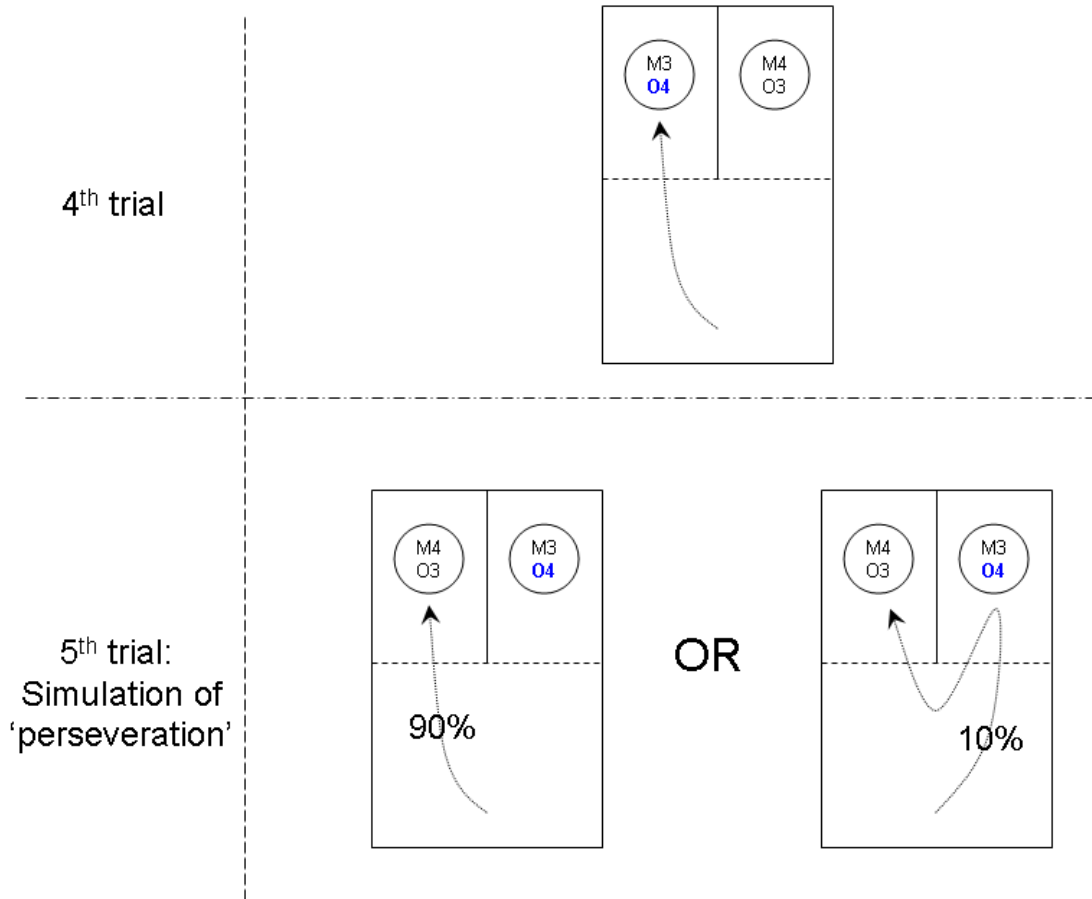


Figure 5.1: an example of generating a simulated observation for a learning trial. The configuration of medium and odour in each bowl for the 5<sup>th</sup> trial and food-associated stimulus (*O4*) were directly obtained from the pre-determined bowl designs as used traditionally in the ID stage of the 7-stage task. If we know that the virtual rat would use spatial perseveration to choose a bowl at the 5<sup>th</sup> trial (Table 5.1), and the rat chose the left bowl at the previous trial (upper figure), then the simulated choice at the 5<sup>th</sup> trial would be the *left bowl* (lower left or lower right figure). Since the food reward is associated with odour *O4* (bold blue) and it is not in the chosen (left) bowl, the rat did not get reward at the 5<sup>th</sup> trial. Because the rat would use spatial perseveration to choose the bowl at the 5<sup>th</sup> trial, the rat would more probably (90%) go to dig in the left bowl directly rather than approach the right bowl and then go to dig in the left bowl. Therefore, the simulated observation would be like the lower left figure with probability 90% and like the lower right figure with probability 10%.

With the above process, one simulation was generated for the sequence of known response patterns and then analysed by our Bayesian model. Figure 5.2 is an example of a Bayesian analysis result, which supports the idea that high (e.g., larger than 0.6) probability of a spatial hypothesis does indicate that the rat may really be using the corresponding spatial response pattern to choose bowls.

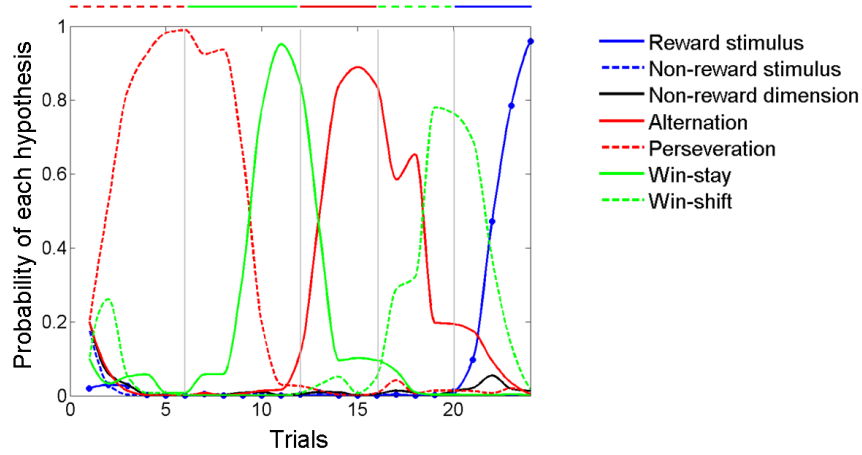


Figure 5.2: Bayesian probability of each hypothesis over trials of a simulation when the virtual rat used spatial perseveration (dotted red), win-stay (solid green), spatial alternation (solid red), and win-shift (dotted green) response patterns to choose bowls. This is just one of the 60 stimulations. Solid blue curve: the posterior probability of the reward-associated perceptual hypothesis over trials; dotted blue curve: the posterior probability of non-reward associated perceptual hypothesis in the relevant dimension. Black curve: the sum probability of the two perceptual hypotheses in the irrelevant dimension. The red and green curves represent the posterior probabilities of the four spatial hypotheses over trials. Each grey vertical line demonstrates the onset of a new ground-truth response pattern (see Table 5.1), and each colour bar on the top represents the ground-truth response pattern for a set of consecutive trials.

The Bayesian analysis result from one simulation may not be representative enough due to stochastic simulation of 1<sup>st</sup>/2<sup>nd</sup> observation for each trial. Here, 60 simulations were generated with the same above process and the same series of known hypotheses (Table 5.1). The only possible difference is the 1<sup>st</sup>/2<sup>nd</sup> observation for each trial between stages. Over the 60 simulated stages, the highest Bayesian probability of each spatial hypothesis from each stage was collected, resulting in the histogram of the maximum probability of each hypothesis over the 60 stages (Figure 5.3). Figure 5.3 shows that the highest probability of each spatial hypothesis in most simulations is larger than 0.6, particularly for spatial perseveration, spatial alternation, and win-stay response patterns. This result confirms that high Bayesian probability of a spatial hypothesis often indicates that the simulated rat did use the corresponding spatial response pattern to choose bowls. Thus, we believe that the use of the Bayesian approach on rats performing the experiment is likely to capture *bona fide* regularities in their behaviour.



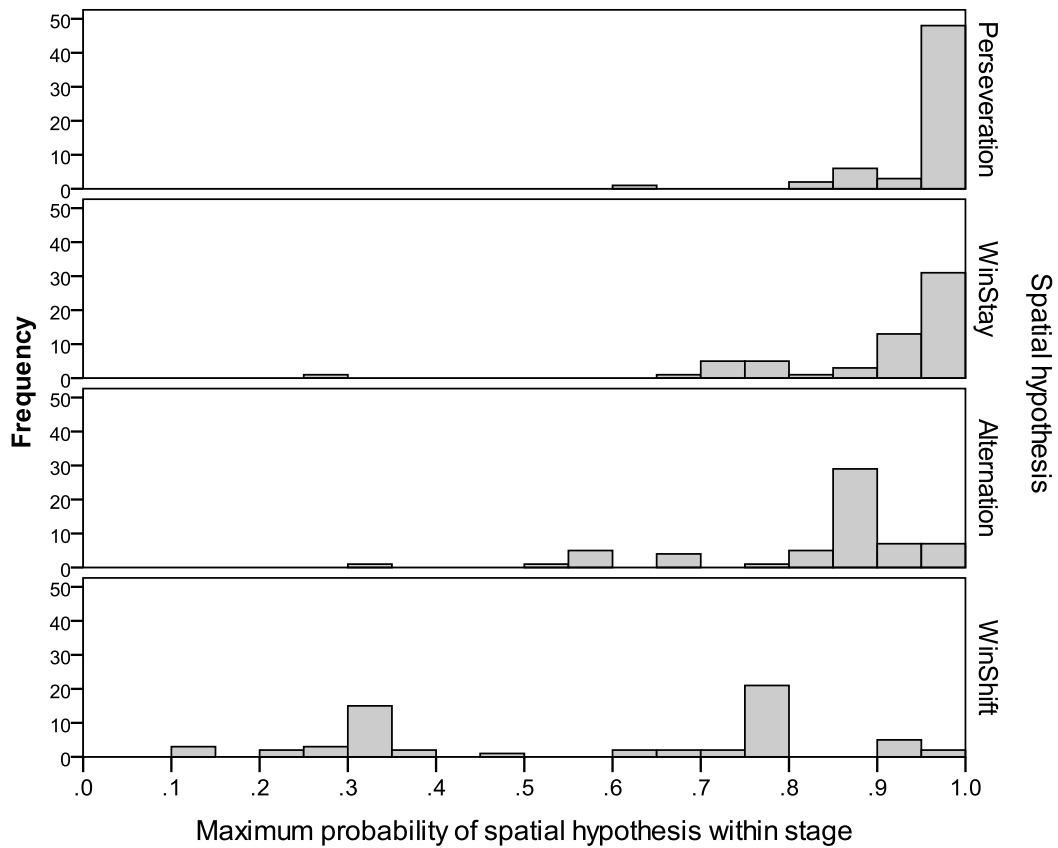


Figure 5.3: distribution of each spatial hypothesis' highest probabilities over simulated stages. The highest probability of each spatial hypothesis in most simulated stages is larger than 0.6, particularly for spatial perseveration, spatial alternation, and win-stay response patterns.

Interestingly, Figure 5.3 (last histogram) also shows that the Bayesian estimate of the win-shift response pattern is quite low (even less than 0.3) in some simulations, even though the virtual rat did use this spatial response pattern. Figure 5.4 demonstrates an example simulation within which the highest probability of the hypothesis associated with win-shift response pattern (dotted green curve) is  $\sim 0.3$ . Detailed inspection of the simulated stage's observation found that the virtual rat selected the 2<sup>nd</sup> bowl at trial 19 where the spatial win-shift response pattern is simulated, leading to low likelihood of the hypothesis associated with the win-shift response pattern, which in turn led to the low Bayesian estimate of the win-shift.

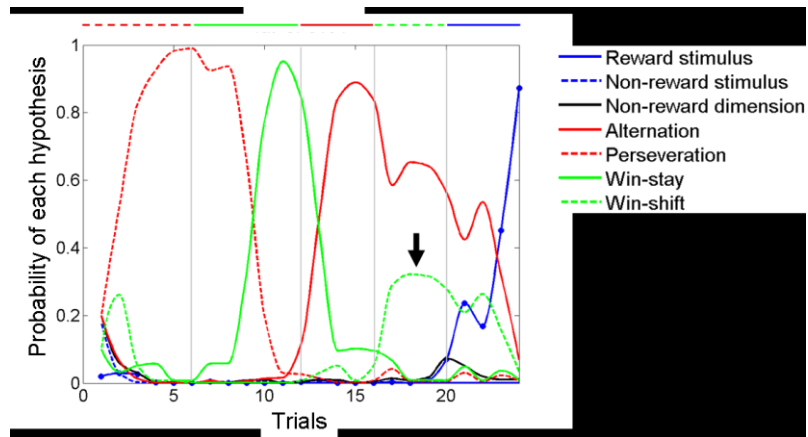


Figure 5.4: Bayesian probability of each hypothesis over trials of a simulation when the virtual rat did try all four spatial hypotheses. The Bayesian estimate of the hypothesis associated with win-shift response pattern is not high enough (black arrow) although the virtual rat did use it for bowl choice (between the two grey vertical lines). Each colour bar on the top represents the ground-truth response pattern for the corresponding set of consecutive trials.

The above analysis focused on spatial response patterns. However, the irrelevant perceptual dimension is the interest of focus at the beginning of the ED stage. One similar doubt is how often a high Bayesian probability of the irrelevant dimension really indicates that the rat used the stimuli of the irrelevant dimension to dig in bowls. With the same simulation process as above, one simulated ED stage was generated based on another series of known hypotheses (Table 5.2). The Bayesian analysis of the simulation data (Figure 5.5) supports the idea that high Bayesian probability of the irrelevant dimension (black curve; the sum of the probability of the two perceptual hypotheses in the irrelevant perceptual dimension) does indicate that the rat may really be using a certain perceptual response pattern of the irrelevant dimension to dig in bowls.

Table 5.2: a series of known response patterns used to generate simulated ED stage data. To generate a simulated ED stage, one of the two perceptual hypotheses from the irrelevant dimension is randomly chosen once and used for the first four trials. Although the simulated ED stages contain some spatial response patterns, here we focus our analysis on the first several trials, as the beginning of the ED stage is one area of focus in the analysis of real rats' data.

Trial	1-4	5-7	8-11	12-17
Response pattern	One perceptual hypothesis from the irrelevant perceptual dimension	Spatial perseveration	Win-shift	Rewarded stimulus

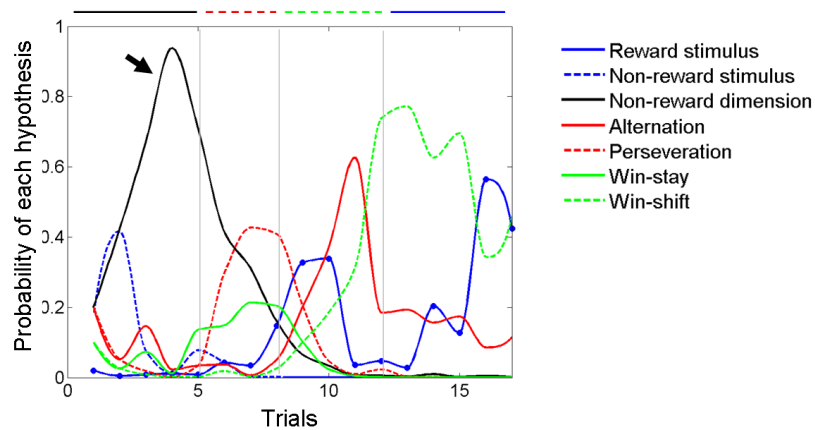


Figure 5.5: Bayesian probability of each hypothesis over trials of one simulated ED stage (out of 60 simulated stages) when the virtual rat used one stimulus of the irrelevant dimension to choose bowls in the first four trials. The Bayesian analysis did find the rat used certain stimuli of the irrelevant dimension to choose bowls at the beginning of the stage (arrow).

Similar to the analysis for spatial hypotheses, 60 simulation stages were generated based on the series of known response patterns (Table 5.2). Over the 60 simulations, the highest Bayesian probability of the irrelevant dimension from the first four trials of each simulated ED stage was collected, resulting in the histogram of the maximum probability of the irrelevant dimension over simulations (Figure 5.6). Again, Figure 5.6 supports the idea that high Bayesian probability (e.g., larger than 0.6) of the irrelevant dimension often indicates that the rat did use a certain perceptual response pattern of the irrelevant dimension to choose bowls.

However, Figure 5.6 also shows that the Bayesian estimate of the irrelevant dimension is just around 0.2 in a few simulated stages. Figure 5.7 demonstrates an example of such a simulation. Although the virtual rat actually used one perceptual stimulus of the irrelevant dimension to choose bowls over the first four trials, detailed inspection of the example stage's observation showed that the virtual rat selected the 1<sup>st</sup> bowl and the bowl choice side alternated over the first four trials, leading to the high Bayesian estimate of the hypothesis associated with the spatial alternation response pattern, and therefore, low Bayesian estimate of irrelevant dimension. In other words, the simulated rat's choices in the first four trials support multiple

response patterns, and the Bayesian analysis seems to reflect genuine ambiguity in the response patterns.

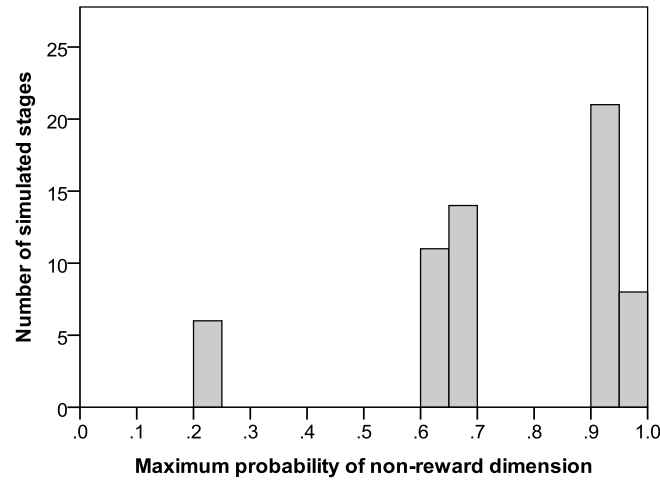


Figure 5.6: distribution of highest probabilities of the irrelevant dimension from the first four trials over simulation stages. The highest probability of the irrelevant dimension in most stages is larger than 0.6. The 1<sup>st</sup>/2<sup>nd</sup> observation is randomly generated for each of the first four trials, leading to multiple different values of the highest probability of the irrelevant dimension for different simulation data.

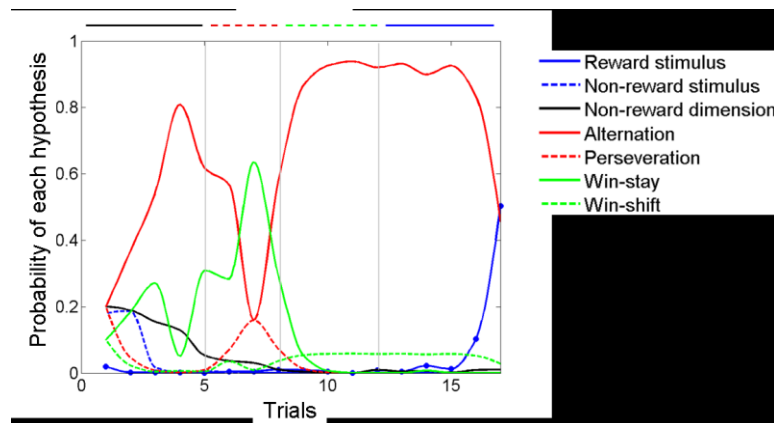


Figure 5.7: Bayesian probability of each hypothesis over trials of one simulation when the virtual rat used one perceptual stimulus of irrelevant dimension to choose bowls in the first four trials, but the Bayesian estimate of the irrelevant dimension is very low (around 0.2). There are six simulations having similar low Bayesian estimates.

## 5.2 Bayesian analysis of spatial patterns in simulated data

The number of spatial trials within a certain stage (e.g., ED stage) has been estimated and compared between two groups of rats based on Bayesian analysis of real data (see Section 3 for details), where *spatial trials* are those trials in which the rat used certain spatial response patterns to choose bowls. The findings about spatial trials on the real data are convincing only if the Bayesian estimates of spatial hypotheses did reflect the

actual spatial response patterns used by rats. In other words, higher estimated numbers of spatial trials within a stage should really indicate that rats actually spend more trials using corresponding spatial response patterns to choose bowls. To confirm the effectiveness of the Bayesian analysis in estimating spatial response patterns, simulation data with known spatial response patterns were generated and then analysed by the Bayesian approach as below.

First, 160 simulated rats' ID stage data were generated, with each stage consisting of a maximum of 20 trials. The data were generated by implementing the following procedure in MATLAB:

For each spatial response pattern  $h$  from the set of the four spatial response patterns:

- For each consecutive trial number  $k$  from the set  $\{3, 5, 7, 9\}$ 
  - Repeat the following procedure 10 times, to generate 10 simulated ID stages
    - ◆ Generate a sequence of 20 ground-truth response patterns, with an example of the generation process shown in Figure 5.8 where the response pattern  $h$  is *perseveration* and  $k=5$ . The generation process is under two constraints, (1) the last six response patterns being reward-associated and (2)  $k$  consecutive trials with the spatial response pattern  $h$  appearing randomly in the sequence (but not in the last six trials). All the other positions in the sequence were filled randomly with other response patterns (including the non-rewarded and two irrelevant perceptual and the other three spatial response patterns), under the constraint that any chosen response pattern was filled in 2 to 4 consecutive trials if possible.
    - ◆ For the just-generated sequence of 20 ground-truth response patterns, generates the rat's simulated behaviour (side choice, got reward or not, 1<sup>st</sup>/2<sup>nd</sup>) for each of the 20 trials as described in Section 5.1 (Figure 5.1).

Step 1	Trial	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15-20
	Response pattern															M1
Step 2	Trial	1	2	3	4	5	6	7	8-12				13	14	15-20	
	Response pattern								Perseveration						M1	
Step 3	Trial	1-3		4	5	6	7	8-12				13	14	15-20		
	Response pattern	Alternation						Perseveration						M1		
Step 4	Trial	1-3		4-5		6	7	8-12				13	14	15-20		
	Response pattern	Alternation		O2				Perseveration						M1		
Step 5	Trial	1-3		4-5		6-7		8-12				13	14	15-20		
	Response pattern	Alternation		O2		Win-shift		Perseveration						M1		
Step 6	Trial	1-3		4-5		6-7		8-12				13-14		15-20		
	Response pattern	Alternation		O2		Win-shift		Perseveration				O1		M1		

Figure 5.8: an example demonstrating how to generate a sequence of 20 responses patterns with *perseveration* in five consecutive trials ( $h$  is *perseveration*,  $k=5$ ). Suppose the rewarded response pattern is *M1*, and non-rewarded and two irrelevant perceptual response patterns are *M2*, *O1*, and *O2* respectively. Step 1: set the rewarded response pattern *M1* to the last six trials. Step 2: randomly select five consecutive trials (except for the last six trials) and set the *perseveration* response pattern to them. Steps 3-6: randomly select one response pattern from the remaining 3 perceptual and 3 spatial response patterns, and set it to consecutive trials of length 2-4 every time. Red trials: not yet filled with response patterns so far; dark green trials: being filled with a response pattern; light green trials: already filled with response patterns.

Once the 160 virtual rats' ID stage data were generated from the above procedure, each virtual ID was then analysed using the Bayesian approach. For each virtual ID that contains a ground-truth number of trials of specific spatial response pattern  $h$ , the estimated number of trials of this spatial response  $h$  was collected based on the corresponding Bayesian estimates (as done for real rat data in Section 3).

Polynomial (linear) contrast was conducted between the estimated numbers of trials of the spatial hypothesis ('estimated spatial trials' in short; dependent variable) and the ground-truth number of trials of spatial hypothesis ('ground-truth spatial trials' in short; independent variable) (Figure 5.9). Ground-truth spatial trials had a significant

effect on estimated spatial trials,  $F(1, 156) = 101.45$ ,  $p < 0.001$ , with the more trials in which rats used spatial response patterns to choose bowls, the more estimated spatial trials were from the proposed Bayesian approach. Figure 5.9 also shows that the mean estimated spatial trials were always less than the ground-truth spatial trials by about 1.5 to 2.5 trials. This is reasonable because it often takes a few trials for the posterior probability of the spatial hypothesis to increase to a high value, such that the estimated number of spatial trials is often smaller than the actual number of spatial trials.

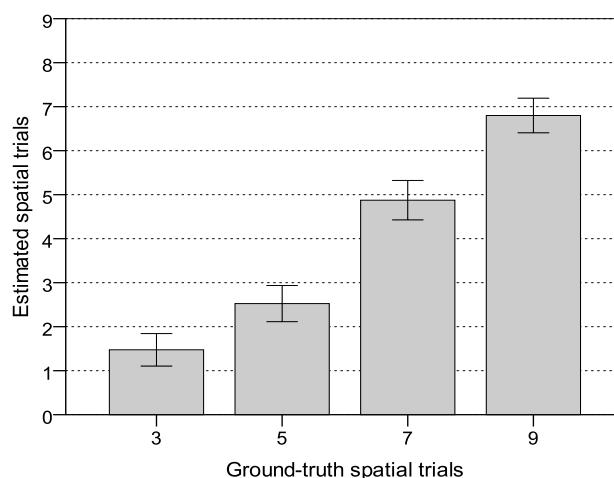


Figure 5.9: the estimated number of spatial trials increases with more ground-truth spatial trials. Vertical line segments represent the standard errors. For each virtual ID that contains a ground-truth number of trials of specific spatial response pattern  $h$ , the estimated number of spatial trials was collected where Bayesian estimate of the spatial response pattern  $h$  is larger than 0.6.

### 5.3 Evaluation of Bayesian criterion in simulated data

The Bayesian approach provides an alternative learning criterion to determine whether or not a rat has learned the reward-associated stimulus in each learning stage. More specifically, if the posterior probability of the reward-associated hypothesis is larger than a pre-defined high threshold at a certain trial (e.g., 0.95, meaning that we are 95% confident the rat is using the reward-associated stimulus to choose bowls), we would consider that the rat has learned the stimulus-reward association. Note that the high threshold 0.95 is only compared to the posterior probability of the *reward-associated* hypothesis in the Bayesian learning criterion, while another threshold 0.6 used in the thesis is to determine whether *non-reward associated* response patterns

were used before learning the reward-associated response pattern in each stage.

To show whether the Bayesian criterion works appropriately, 40 simulation stages were generated, with each stage 20 trials long. More specifically, to generate a simulation stage, a series of 20 ground-truth response patterns were manually pre-determined such that the following bowl choice results (Table 5.3) were generated based on the predetermined real bowl designs of the ID stage. The only possible differences among the 40 simulated data were the 1<sup>st</sup>/2<sup>nd</sup> observation in trials.

Table 5.3: the sequence of 20 bowl choice results for each of the 40 simulations.

Trial	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15-20
Rewarded	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	<div> <div>✓</div> <div>✓</div> <div>✓</div> <div>✓</div> <div>✓</div> <div>✓</div> </div>

The 40 simulated stages were then analysed by the Bayesian approach, and the first trial to Bayesian criterion (where posterior probability of the reward-associated hypothesis reaches 0.95) was obtained for each simulated stage. Figure 5.10 shows that, with Bayesian criterion, the virtual rats would be considered to have learned the stimulus-reward association around trial 10 to 12 in most of the 40 simulated stages (also see Figure 5.11, left). In comparison, the rat has to take the full 20 trials to finish each stage if using the frequentist 6-in-a-row criterion. Interestingly, from Table 5.3, it did look like the virtual rat has learned around trial 10 or 11.



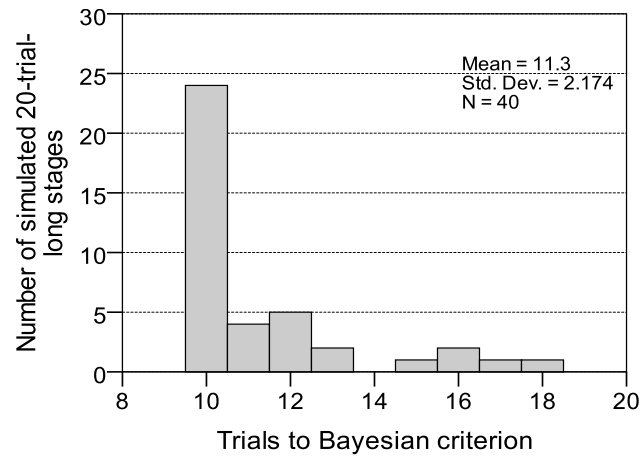


Figure 5.10: histogram of trials to Bayesian criterion (with threshold 0.95) over 40 simulations, with each stage 20 trials long and having the same bowl choice results as seen in Table 5.4, but often having different sequence of 1<sup>st</sup>/2<sup>nd</sup> observations between simulations.

One observation from Figure 5.10 is that the trials to Bayesian criterion are more than 15 in a few stages, although all the simulated stages have the same sequence of bowl choice results. Detailed investigation showed that the longer trial to Bayesian criterion comes from the dominance of a spatial response pattern *win-shift* around trial 10 (Figure 5.11, right), the Bayesian estimate of which increases to a high value (0.8) mainly due to the consistency between bowl choices and the spatial hypothesis associated with the *win-shift* response pattern in trials 6 to 10. In comparison, the correct hypothesis is not consistent with the rat's bowl choice in trial 8, leading to the decrease in the probability of the correct hypothesis from trial 8. Again, this supports the idea that the Bayesian approach handles true ambiguity well because it considers a set of hypotheses rather than just the frequentist null hypothesis.

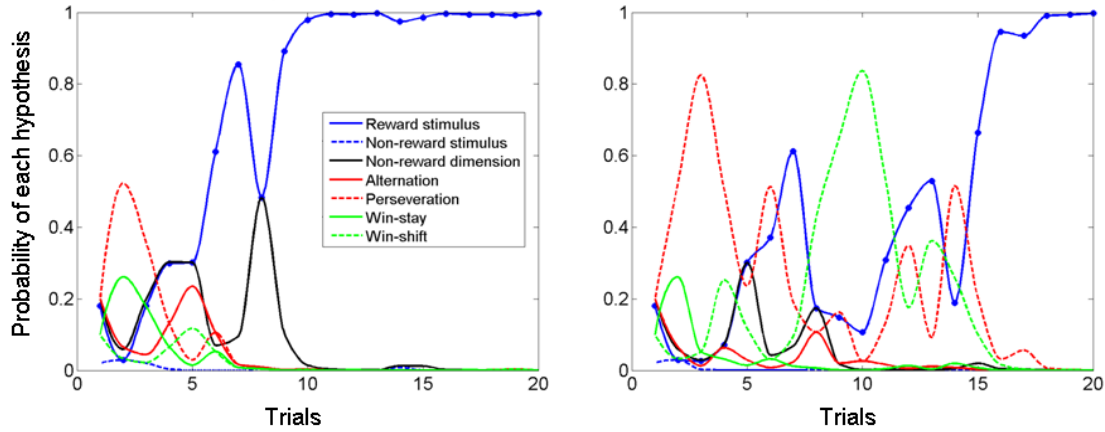


Figure 5.11: probability of each hypothesis over trials for two simulated stages. Left: the estimated trial to Bayesian criterion is 10; Right: the estimated trials to Bayesian criterion are 18, probably because the spatial hypothesis associated with the win-shift response pattern (dotted green curve) is consistent with a few trials' observation where the virtual rat continuously chose the 1<sup>st</sup> bowl without going to check the 2<sup>nd</sup> bowl.

The above result indicates that Bayesian criterion is consistent with the frequentist 10-correct-out-12 criterion (see the 10-correct-out-of-12 trials in Table 5.3). To show that the Bayesian criterion is consistent with multiple different frequentist learning criteria, we also applied Bayesian analysis to another 40 simulations, with each stage just six trials long and having the same bowl choice result (Table 5.4). All the 40 simulation stages were different only in the 1<sup>st</sup>/2<sup>nd</sup> observation which was randomly generated for each trial in each simulation.

Table 5.4: the same sequence of six bowl choice results for each of the 40 simulations.

Trial	1	2	3	4	5	6
Rewarded	✓	✓	✓	✓	✓	✓

Figure 5.12 shows that the Bayesian criterion (with threshold 0.95) is satisfied or close to satisfied at the last trial in most of the 40 simulations, supporting the idea that the Bayesian criterion is consistent with the traditional 6-in-a-row criterion at least in this simple condition (where a stage consists of just six trials).

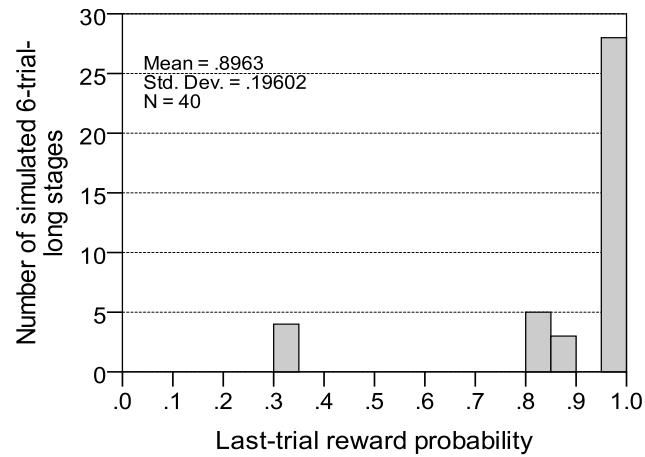


Figure 5.12: histogram of Bayesian probability of the reward-associated hypothesis in the last trials of the 40 simulations, with each stage just six trials long. Bayesian criterion is satisfied or close to be satisfied in the last trials of most simulations.

For the four simulated stages, for which the last-trial reward probability is around 0.35 (Figure 5.12), detailed investigation told us that the observation from trials 2 to 5 is consistent with the *spatial alternation* response pattern (particularly choosing the 1<sup>st</sup> bowl without checking the 2<sup>nd</sup> bowl), causing the higher Bayesian probability of the corresponding hypothesis (Figure 5.13 right, solid red curve). Interestingly, in this case, we may also think the rat used *spatial alternation* to choose bowls. Therefore, the lower Bayesian estimate of reward hypothesis in the last trial, while not consistent with 6-in-a-row criterion, seems to be consistent with a human's subjective judgement. Note that if such a pattern were seen in a real rat, it would be considered as having learned the stage by the 6-in-a-row criterion, whereas the closer level of analysis by the Bayesian approach allows highlighting of the ambiguity.

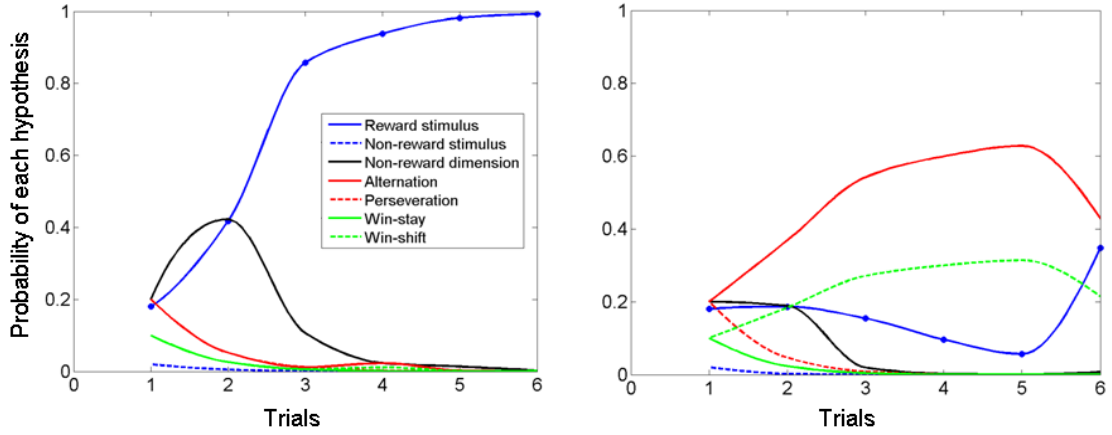


Figure 5.13: probability of each hypothesis over trials for two simulations. Left: an example simulation stage in which Bayesian criterion is consistent with the 6-in-a-row criterion; Right: one simulation with Bayesian criterion not satisfied at last trial, due to the higher Bayesian estimate of the hypothesis associated with the *spatial alternation* response pattern which is consistent with the observed data from trials 2 to 5.

The above two sets of results together indicate that Bayesian criterion may summarise behavioural choice well in multiple conditions in which rats may have learned the stimulus-reward association and seems quite consistent with human subjective judgement on rats' learning. Therefore, the Bayesian criterion may be an appropriate alternative to the frequentist 6-in-a-row criterion when determining whether a rat has learned the stimulus-reward association. In addition, the Bayesian approach handles true ambiguity of response patterns well, which occasionally occurs even if there has been careful counterbalancing of conditions, because it depends in part on how the rat responds.

#### 5.4 Justification of the likelihood function

In the proposed Bayesian model, the 1<sup>st</sup>/2<sup>nd</sup> observation at each trial is included in the likelihood function in order to differentiate the likelihood of spatial hypotheses from perceptual hypotheses. Here we use the simulated data to provide more *evidence for the effect of the 1<sup>st</sup>/2<sup>nd</sup> observation in the likelihood* on the correct estimate of response patterns.

Just as described in Section 5.1, we generated one *representative* simulation based on

a sequence of known response patterns (Table 5.5; reproducing Table 5.1) and then analysed the simulated stage data using (1) the proposed Bayesian model (i.e., with 1<sup>st</sup>/2<sup>nd</sup> observation *included* in the likelihood function) and (2) the simplified Bayesian model (i.e., with ‘1<sup>st</sup>/2<sup>nd</sup>’ observation *excluded* in the likelihood function).

Table 5.5: a series of ground-truth response patterns used to generate simulation data. ‘Random’ means the virtual rat would randomly dig either left or right bowl.

Trials	1	2-5	6-11	12-15	16-19	20-25
Hypothesis	Random	Spatial perseveration	Win-stay	Spatial alternation	Win-shift	Rewarded stimulus

The analysis did show that the 1<sup>st</sup>/2<sup>nd</sup> observation in the likelihood function is very useful to estimate correctly what response patterns were being used by rats. Specifically, Figure 5.14 (right) shows that, with the simplified Bayesian model which excluded the 1<sup>st</sup>/2<sup>nd</sup> observation in the likelihood function, the *spatial alternation* response pattern was not found, while a certain erroneous perceptual response pattern (black curve) was estimated. In comparison, the proposed Bayesian model with 1<sup>st</sup>/2<sup>nd</sup> observation included in the likelihood function correctly estimated all the spatial response patterns (Figure 5.14, left). This result supports the idea that 1<sup>st</sup>/2<sup>nd</sup> observation should be included in the likelihood function in the Bayesian model.

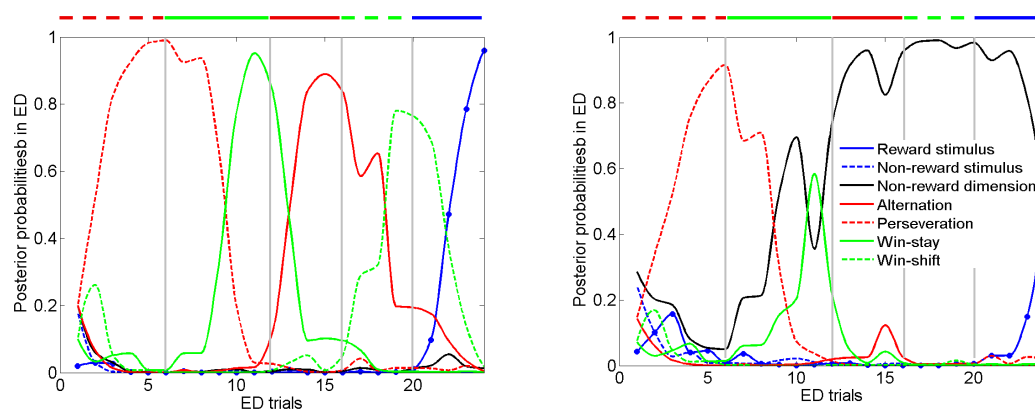


Figure 5.14: representative Bayesian analysis results with two different Bayesian models. The Bayesian model considering 1<sup>st</sup>/2<sup>nd</sup> observation (left) performs better than the Bayesian model without 1<sup>st</sup>/2<sup>nd</sup> observation (right). The colour bars above each figure represent the ground-truth response patterns for each set of consecutive trials.

All the above Bayesian analyses on the simulated data provide convincing evidence

for the validity of Bayesian analysis on rats' data. In particular, high Bayesian probability of a hypothesis often indicates that the rat used the corresponding response patterns to dig bowls, and higher estimated number of spatial trials within a stage really indicates that rats actually use spatial response patterns to choose bowls in more trials. In addition, the appropriateness of the Bayesian learning criterion and the effect of 1<sup>st</sup>/2<sup>nd</sup> observation in the likelihood function were also supported by the Bayesian analysis on simulated data.

## **Chapter III: Bayesian Analysis of Human Set-shifting**

While Bayesian analysis has been shown to be promising in analysing the 7-stage task in rats, we evaluated the validity of the Bayesian analysis on the basis of simulated data. To further explore the validity of the Bayesian analysis, I used an analogous human 7-stage task, similar to the CANTAB<sup>®</sup> ID/ED task (Roberts et al., 1988), to collect human participants' oral reporting of the response strategy that they used to make their choice in each trial. By correlating the participants' subjective reports and the Bayesian estimates of response patterns, I hoped to evaluate whether the Bayesian estimates match the response intentions of human participants.

Given the apparent differences in the way in which humans and rats approach the task, I found that the Bayesian model developed for rats could not be directly applied to the human task. Therefore, I modified the approach that I used in rats to consider the effect of feedback consequent to each response choice in the task (reward vs nonrewarded) on learning. Adding this information for human participants improved the applicability of the Bayesian analysis, but it did not do so for rats. This implies that humans might engage in hypothesis-testing, using feedback to eliminate false hypotheses about the correct choice, whereas rats might use simple associative mechanisms in which the effect of feedback is subtler on a trial-by-trial basis.

### **6 Method for human attentional set-shifting**

#### **6.1 Participants and apparatus**

Thirty-eight students from the University of St Andrews were recruited, mainly from the University's online SONA system, comprising twenty-six female and twelve male participants and with an age range of 18-25.

A task (Figure 6.1) analogous to the CANTAB IDED task (Roberts et al., 1988) was coded in MATLAB (The MathWorks, Inc.) and run on a common modern 14-inch

Lenova laptop (IdeaPad 510S) in quiet testing rooms. The software was displayed at the centre of the laptop screen, with the size of the software graphical user interface (GUI) approximately 640-by-620 pixels, and with normal viewing distance (around 20-25 inches) between the screen and the participant. A box containing the combination of a blob stimulus and a stick stimulus (or just one of them for the SD stage) was displayed on the top half of the task window (Figure 6.1). Participants could click the button immediately below the box to change back and forth between the first and second stimulus (or stimuli). The task used stimuli and stage structure analogous to the CANTAB ED/ED but was modified to make it more functionally similar to the rat 7-stage task, allowing the participant to see stimuli in only one box at a time. The participants could make their choice simply by clicking on the box framing the blob and stick stimuli and indicated their confidence in their choice by moving a sliding pointer along a horizontal scroll bar. The microphone built into the laptop was used to collect each participant's oral report of the reason for the choice they made in each learning trial.

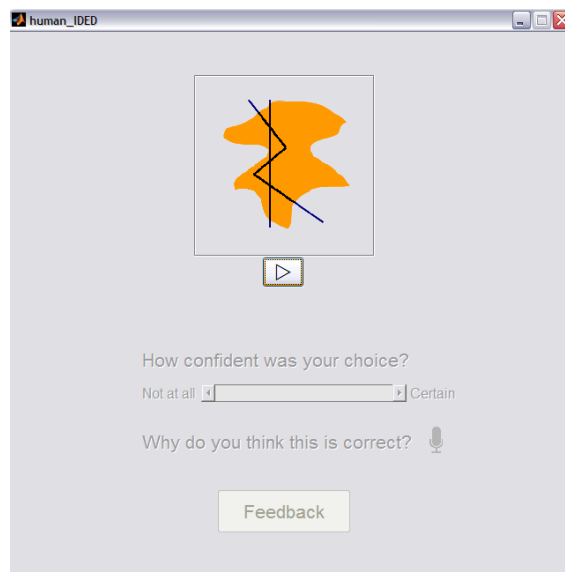


Figure 6.1: a screen-shot of the 7-stage task software for an example learning trial. In the window, the top box contains a pair of orange blob and black stick(s) stimuli. Participant can click the arrow button below the box to view another pair of stimuli.

## 6.2 Materials and procedure

The human 7-stage task, like the rat 7-stage task, consists of a simple discrimination learning stage and six compound discrimination learning stages in the order simple



discrimination (SD) – compound discrimination (CD) – reversal (REV1) – intra-dimensional (ID) – reversal (REV2) – extra-dimensional shift (ED) – reversal (REV3). The trial design for each learning stage is the same as that for the rat task, with the exception that the two dimensions of stimuli in the human task are ‘blob’ and ‘stick’ (Figure 6.2) rather than ‘medium’ and ‘odour’ as in the rat task.

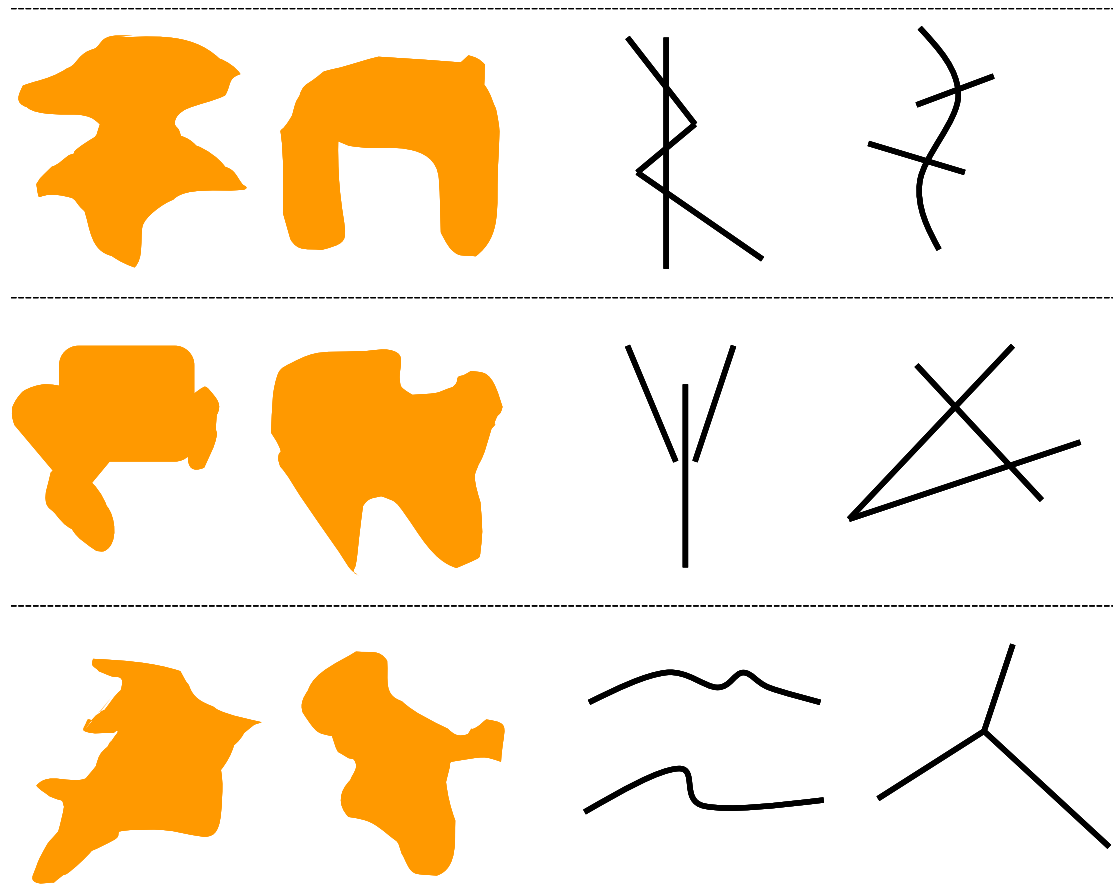


Figure 6.2: stimuli used in the human 7-stage task. Two dimensions of stimuli are used: blob (first two columns) and stick (third and fourth columns). The stimuli in each row are used for a different compound stage and subsequent reversal stage.

Before starting the task, each participant was asked to watch a short video on the computer screen demonstrating how to make a choice and how to answer the questions about their choice confidence and choice reason. Each participant was told to try to make a correct choice at each trial. The program was then initiated on the laptop to begin testing. Each participant was first required to input his or her gender and age (just for statistical information of participants) on the software GUI and then

click a 'Start' button to initiate the first trial of the first learning stage in the 7-stage task.

In each trial, the participant could view one or the other pair of stimuli simply by clicking the button below the stimuli box (Figure 6.3, top row). The participant could choose one of the paired stimuli by clicking on the box containing the pair; The participant could then use a mouse to drag the scroll bar to a position on the slider representing their confidence level for the current choice (Figure 6.3, middle left). Then the participant should answer the second question 'Why do you think this (i.e., your choice) is correct?' orally and the software automatically recorded their voice using the built-in microphone and saved the recording into a new and unique sound file for each trial (Figure 6.3, middle right). These files were later translated into digital text files with assistance from Meggie Rix, who volunteered to translate the oral reports. Once the voice was recorded, the participant could press the 'Choice result' button to see the choice result displayed on the right side of the top box (Figure 6.3, bottom left), and then click the same button with the changed label 'Next trial' to advance to the next trial (Figure 6.3, bottom right). The above process was repeated in each learning trial until the participant made six correct choices in a row. A new stage was then started for the participant to learn. The participant could, however, stop the task at any time if they did not feel comfortable or for any other reason.

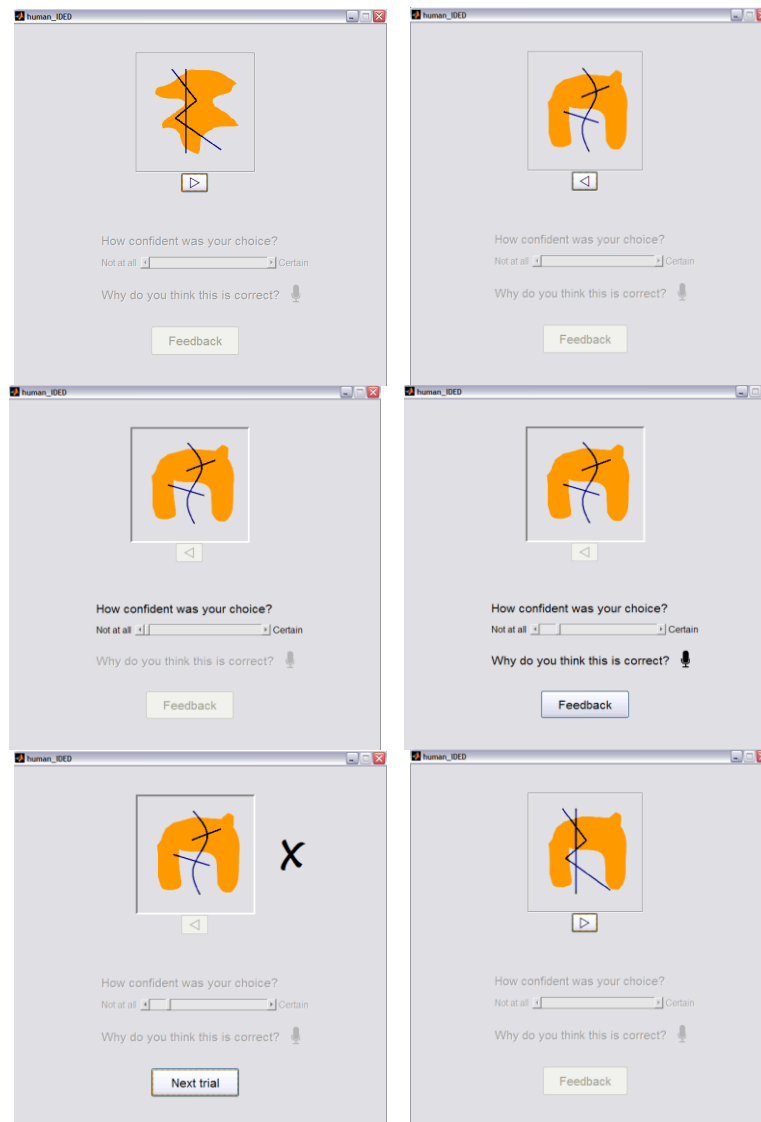


Figure 6.3: screen-shots of the software to demonstrate the procedure when performing the task. Top two: before making a choice, the participant could press the arrow button below the stimuli box to view either the first pair of stimuli (top-left) or the second pair of stimuli (top-right). Middle left: the participant chose the second pair of stimuli by clicking the box containing the stimuli. Middle right: the participant dragged the bar on the slider to a position representing their confidence that the choice was correct; this activated the microphone icon, which prompted the participant to verbalise the reason for their choice. Bottom left: after pressing the 'Feedback' button, the choice result appeared on the right side of the top box, and the 'Feedback' button became 'Next trial' button. Bottom right: the participant clicked the 'Next trial' button and a new trial started, in which the combination of blob and stick could be changed.

Typically, participants finished the task in around 40 minutes and were then debriefed about the true purpose of the study. This study received ethical approval from the School of Psychology and Neuroscience Ethics Committee (SEC), with approval code PS12391 (Appendix A). Relevant forms for advertisement, debriefing, participant

information, etc. are also appended (Appendices B- F).

### 6.3 Human behavioural simple Bayesian model

We used Bayesian analysis to quantify the rats' performance using the classical Bayes' rule (Equation 2.1 in Section 2):

$$P(H = h_i | \{E, \mathbf{E}^*\}) = \frac{P(E | \{H = h_i, \mathbf{E}^*\}) \cdot P(H = h_i | \mathbf{E}^*)}{P(E | \mathbf{E}^*)} \quad (6.1)$$

Note that the comma ',' in the likelihood  $P(E | \{H = h_i, \mathbf{E}^*\})$  means 'and', therefore  $P(E | \{H = h_i, \mathbf{E}^*\})$  represents the probability of E given  $H = h_i$  and  $\mathbf{E}^*$ . The performance of participants in the human attentional set-shifting task may also be analysed by the behavioural Bayesian model that I developed for the rat task. However, considering the differences between humans and rats in performing the tasks, the behavioural Bayesian model needs to be modified from the following aspects before being applied to the human task.

#### Likelihood

The likelihood item about the '1st/2nd' observation in the rat behavioural Bayesian model is mainly to differentiate spatial rules (or 'hypotheses') from perceptual rules (or 'hypotheses'). However, human participants cannot have adopted any spatial rules to make their choice because stimuli were displayed not on the left and right but on the centre in the human task. So the likelihood item about the '1st/2nd' observation is removed, and the likelihood function is simply designed as

$$P(E | H = h_i, \mathbf{E}^*) = \begin{cases} \alpha & \text{if } h_i \text{ consistent with observation } E \\ 1 - \alpha & \text{if } h_i \text{ inconsistent with observation } E \end{cases} \quad (6.2)$$

This likelihood function is exactly the same as that developed for my Honours-year project and can be considered as a simplification of the one developed in Section 2. As discussed before,  $\alpha$  should be set to a high value close to 100%. I used  $\alpha = 0.85$ .

#### Hypotheses

Since the human participants could not use spatial rules to make their choices, the

four spatial hypotheses used for the rat behavioural Bayesian model were removed. Instead, I observed, from the participants' oral reports, that sometimes they used the combination of blob and stick to make their choice, and sometimes they used the alternation of two blobs or alternation of two sticks over trials to make choices (Table 6.1). Therefore, I added four more 'combination' hypotheses and two stimuli alternation hypotheses, resulting in the following 10 perceptual hypotheses  $H$ :

- Hypothesis 1 ( $h_1$ ): participant uses the blob B1 to make choice.
- Hypothesis 2 ( $h_2$ ): participant uses the other blob B2 to make choice.
- Hypothesis 3 ( $h_3$ ): participant uses the stick L1 to make choice.
- Hypothesis 4 ( $h_4$ ): participant uses the other stick L2 to make choice.
- Hypothesis 5 ( $h_5$ ): participant uses blob B1 and blob B2 alternatively to make choice.
- Hypothesis 6 ( $h_6$ ): participant uses stick L1 and stick L2 alternatively to make choice.
- Hypothesis 7 ( $h_7$ ): participant uses the combination of blob B1 and stick L1 to make choice.
- Hypothesis 8 ( $h_8$ ): participant uses the combination of blob B1 and stick L2 to make choice.
- Hypothesis 9 ( $h_9$ ): participant uses the combination of blob B2 and stick L1 to make choice.
- Hypothesis 10 ( $h_{10}$ ): participant uses the combination of blob B2 and stick L2 to make choice.

With the exception of the above two aspects, all the details of the simplified Bayesian model are the same as those of the rat behavioural Bayesian model described in Section 2. In the following, I call this simplified Bayesian model the *human*

*behavioural simple Bayesian model.*

Table 6.1: the number of participants using each response pattern with corresponding hypothesis, as determined by their self reports. *B1B2*: blob alternation, *L1L2*: stick alternation; *B1+L1*: combination of B1 and L1, similarly for the other three combinations.

















Hypotheses	<i>B1</i>	<i>B2</i>	<i>L1</i>	<i>L2</i>	<i>B1B2</i>	<i>L1L2</i>	<i>B1+L1</i>	<i>B1+L2</i>	<i>B2+L1</i>	<i>B2+L2</i>
Participants ( <i>n</i> =33)	33	33	33	33	4	6	3	0	3	3











#### 6.4 Human latent probabilistic model and behavioural reward Bayesian model

As well as applying the Bayesian approach to the estimation of participants' behavioural patterns in making choices in the attentional set-shifting tasks, we are also interested in whether the Bayesian approach can be used to infer participants' latent information processing (e.g., latent rules used to make choices) evoked by choice feedback. In the attentional set-shifting task, participants learned the correct rule based on feedback arising as a consequence of their choice (correct/incorrect or rewarded/unrewarded) in each trial within a learning stage. For example, if a participant used the blob stimulus *B1* to make their choice and the choice result is incorrect in one trial, the participant would likely consider that *B1* is not related to the correct choice. In the next trial, the participant would be more likely to make their choice using a different stimulus. On the other hand, if the choice was correct, the participant would be likely to continue using *B1* to make their choice in the next trial. In other words, the correctness of previous choices could largely affect a participant's confidence on which stimuli are more related to correct choice and which are not, and therefore affect their choice in the current trial. Such hypothesis-testing learning appears in human participants and has been supported by participants' oral reports (Table 6.2). Note, however, that rats might have been reliant on associative learning, by which actions and outcomes are learned gradually. In comparison, hypothesis-testing learning is more likely to be driven by high-level cognition (see Section 8 for more discussion). Therefore, in order to estimate our human participants' pattern of

decision-making, the effect of feedback at the end of each trial was incorporated into the Bayesian analysis.

Table 6.2: an example of an oral report from one participant's ED stage, showing the hypothesis-testing learning process. In this example, the participant quickly changed rules if their choice was wrong and continued using the rule if their choice was correct (see last column). The two paired stimuli in each trial are also shown.

Trial	Transcribed self-report and two paired stimuli	Choice feedback	Researcher's remark
	   		
1	<p>I have been presented with new shapes, so I am guessing at this point.</p>  	Correct	Made choice randomly
2	<p>I got the previous question correct and the black line has the same design, so I am going to go with this choice.</p>  	Correct	Made choice based on stick <i>L4</i> because the previous choice included <i>L4</i> and was correct.
3	<p>I got the previous question correct and the black line has the same design, so I am going to stick with it.</p>  	Incorrect	Made choice based on <i>L4</i> because the previous choice included <i>L4</i> and was correct.
4	<p>I am quite surprised that I got the previous question wrong, so maybe I have been wrong about the design of the black line and try this design instead.</p>  	Correct	Choice based on <i>L4</i> was wrong in the previous trial, so made choice based the other stick <i>L3</i> .
5	<p>I got this design of the black line correct in the previous question, so I am just going to stick with the same design.</p>  	Incorrect	Continue using <i>L3</i> because previous choice was correct.
6	<p>I wonder if the orange shape behind plays a role, but at this point I am guessing.</p>  	Correct	Choice based on <i>L3</i> was wrong in the previous trial, so changed rule from stick to blob. Randomly chose one blob <i>B4</i> .

7	I have got a strong feeling that the orange shape behind has got something to do with the correct answer, so I am going to look for the same orange shape.  	Correct	Continue using B4 because previous choice was correct
8	I am quite confident that the orange shape has got something to do with the correct answer, so I am sticking with the same orange shape from the previous question because I got it correct.  	Correct	Continue using B4 because previous choice was correct
9	I am quite sure that the orange shape has got something to do with the correct answer, so I am sticking with the same orange shape from the previous question because I got it correct.  	Correct	Continue using B4 because previous choice was correct
10	I am quite confident of my answer because I have been choosing the same orange shape for the previous three to four questions and I got all of them correct.  	Correct	Continue using B4 because previous choice was correct
11	I am very confident that the orange shape has got something to do with the correct answer.  	Correct	Continue using B4 because previous choice was correct

Before describing the new human model, I would like to clarify the difference between two sets of hypotheses: behavioural hypotheses  $H$  and latent (cognitive) hypotheses  $H'$ . Behavioural hypotheses are the possible hypotheses corresponding to participants' behavioural choice patterns, and have been used in the rat behavioural Bayesian model when analysing rats' behavioural patterns (Section 2). In comparison, latent hypotheses are the possible hypotheses which participants think are *correct*. Behavioural hypotheses  $H$  have been described in the human behavioural simple Bayesian model above (Section 6.3). Latent hypotheses  $H'$  include:

- Hypothesis 1 ( $h'_1$ ), or *B1* hypothesis: participant thinks blob B1 is correct
- Hypothesis 2 ( $h'_2$ ), or *B2* hypothesis: participant thinks blob B2 is correct.
- Hypothesis 3 ( $h'_3$ ), or *L1* hypothesis: participant thinks stick L1 is correct.



- Hypothesis 4 ( $h'_4$ ), or *L2* hypothesis: participant thinks stick L2 is correct.
- Hypothesis 5 ( $h'_5$ ), or *B1B2* hypothesis: participant thinks blob B1 and blob B2 are alternatively correct.
- Hypothesis 6 ( $h'_6$ ), or *L1L2* hypothesis: participant thinks stick L1 and stick L2 are - alternatively correct.
- Hypothesis 7 ( $h'_7$ ), or *B1+L1* hypothesis: participant thinks the combination of blob B1 and stick L1 is correct.
- Hypothesis 8 ( $h'_8$ ), or *B1+L2* hypothesis: participant thinks the combination of blob B1 and stick L2 is correct.
- Hypothesis 9 ( $h'_9$ ), or *B2+L1* hypothesis: participant thinks the combination of blob B2 and stick L1 is correct.
- Hypothesis 10 ( $h'_{10}$ ), or *B2+L2* hypothesis: participant thinks the combination of blob B2 and stick L2 is correct.

Note that each behavioural hypothesis has a *corresponding* latent hypothesis which shares the same stimulus or stimuli, e.g.,  $h_1$  corresponding to  $h'_1$ , and  $h_8$  corresponding to  $h'_8$ . Also note that, in the SD stage, only two stimuli appear in the boxes. Therefore, for SD stages, there are just three latent hypotheses, comprising two hypotheses of single stimulus and one stimulus alternation hypothesis.

Our objective, after the participant made their choice and saw the choice result in the current trial, was to estimate the probability that (the participant thinks) each hypothesis  $h'_i$  was correct, based on the relevant observed data from both the current trial and all previous trials. The relevant observation data in each trial includes (1) the blob and the stick in the chosen box, and the other blob and stick in the un-chosen box, (2) which box was chosen, and (3) whether the choice is correct.

Before formally introducing the new human model, here is an example to show how we, as researchers, could infer an ideal learner's learning process by considering the effect of choice result (Figure 6.4). At the first trial, suppose that the participant looked at the two boxes, chose the box containing Blob 1 and Stick 1, and then found that her choice was wrong. Based on the participant's learning behaviour in this trial, we, as researchers, could infer the following information: Firstly, the participant probably made her choice based on Blob 1 or Stick 1, because she chose the box containing Blob 1 & Stick 1. Secondly, after the participant found her choice was wrong, she would probably exclude Blob 1 and Stick 1 as putative stimuli related to the correct response. This is the effect of choice result on her learning. Then, in the next (2<sup>nd</sup>) trial, suppose that the participant chose the box containing Blob 1 and Stick 2, rather than the box containing Blob 2 and Stick 1, and then found that her choice was correct. Again, we may infer that the participant is more likely to have made her choice based on Stick 2, because she chose the box containing Blob 1 and Stick 2, but after her last trial, she probably thought Blob 1 was not related to the correct response. Also, since the participant found that her choice was correct, she would probably think that Stick 2 was related to the correct response. Again, this is the effect of choice result on her learning. Here we may even predict the participant would be likely to choose the box containing Stick 2 in the next trial. Then, at the third learning trial, we observed that the participant chose the box containing Stick 2. From this observation, we would be more certain that the participant used Stick 2 to make her choice, and also we would think that the participant was more certain that Stick 2 was related to the correct response. Note that this example assumes the participant approaches an 'ideal observer' – namely, somebody that uses the information rationally to solve the problem.

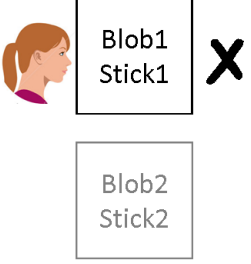
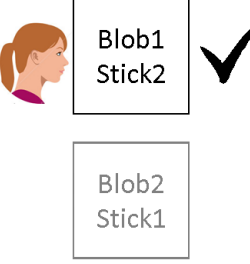
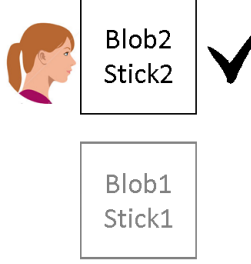
	Trial 1	Trial 2	Trial 3
Participant's learning			
Researcher's Viewpoint (after observing participant's behaviour)	<div style="background-color: #e0ffe0; padding: 5px;">           'You (participant) probably made choice based on Blob1 or Stick1'.         </div> <div style="background-color: #e0f0ff; padding: 5px;">           'You saw your choice was wrong, so probably you would exclude Blob1 nor Stick1 as putative stimuli related to correct response.'            'Effect of choice result'         </div>	<div style="background-color: #e0ffe0; padding: 5px;">           'You made choice more likely based on Stick2.'         </div> <div style="background-color: #e0f0ff; padding: 5px;">           'You saw your choice was right, so probably you would think Stick2 was related to correct response.'            'Effect of choice result'         </div>	<div style="background-color: #e0ffe0; padding: 5px;">           'I am quite sure you used Stick2 to make choice.'         </div> <div style="background-color: #e0f0ff; padding: 5px;">           'I think you are more sure Stick2 is related to correct response.'         </div>

Figure 6.4: example of the inferred effect of choice result on a participant's learning, from the perspective of the researchers. Text in light green is relevant to the behavioural hypotheses, and text in light blue is relevant to the latent (cognitive) hypotheses. Both types of hypotheses are from the perspective of the researchers.

We developed a new Bayesian analysis to estimate the participants' latent deductions. Formally, let  $E$  and  $\mathbf{E}^*$ , respectively, denote the observed data from the current trial and all previous trials. Then all the observed data, from the first trial to the current trial, within a learning stage will be the combined observation  $\mathbf{E} = \{E, \mathbf{E}^*\}$ . To estimate the probability  $P(H' = h'_i | \mathbf{E})$  of a specific latent hypothesis  $h'_i$  after we observed the participant's data  $\mathbf{E}$ , the 'Law of total probability' (Mendenhall et al., 2005) tells us

$$\begin{aligned}
 P(H' = h'_i | \mathbf{E}) &= \sum_{h_j} P(\{H' = h'_i, H = h_j\} | \mathbf{E}) \\
 &= \sum_{h_j} P(H' = h'_i | \{H = h_j, \mathbf{E}\}) \cdot P(H = h_j | \mathbf{E})
 \end{aligned} \tag{6.3}$$

We call Equation (6.3) the *human latent probabilistic model*. In the model,  $P(H = h_j | \mathbf{E})$  is the Bayesian probability of behavioural hypothesis  $h_j$ , which can be estimated in a similar way to that for the rat behavioural Bayesian model.

Psychologically,  $P(H = h_j | \mathbf{E})$  represents how likely it is that the participant used the response pattern (or *rule/strategy* in literature) corresponding to a specific hypothesis  $h_j$  to make their choice at the current trial, after the participant's behaviour (included in  $\mathbf{E}$ ) has been observed. The other term,  $P(H' = h_i' | \{H = h_j, \mathbf{E}\})$ , is the conditional probability of  $h_i'$  given a specific behavioural hypothesis  $h_j$ . This term can be psychologically interpreted as, if the participant used the response pattern corresponding to the specific hypothesis  $h_j$  to make their choice and saw the choice result (included in  $\mathbf{E}$ ) at the current trial, how likely is it that the participant would think the specific latent hypothesis  $h_i'$  is related to the correct response. The product  $P(H' = h_i' | \{H = h_j, \mathbf{E}\}) \cdot P(H = h_j | \mathbf{E})$  of the two terms is summed over all possible behavioural hypotheses because we are not sure which response pattern the participant is using to make their choice in the current trial, but can only estimate the probability of each behavioural hypothesis corresponding to each response pattern.

To estimate the posterior probability  $P(H' = h_i' | \mathbf{E})$  of a specific associative hypothesis  $h_i'$  at the current trial, we need to compute the conditional probability  $P(H' = h_i' | \{H = h_j, \mathbf{E}\})$  of the specific latent hypothesis  $h_i'$  and the posterior probability  $P(H = h_j | \mathbf{E})$  of each behavioural hypothesis  $h_j$ .

### **Feedback effect function $P(H' = h_i' | \{H = h_j, \mathbf{E}\})$**

This conditional probability function is designed to reflect the effect of choice result in the current trial (not in the previous trial) on a participant's confidence in the specific latent hypothesis  $h_i'$  after the participant observed the feedback at the end of the current trial, therefore called 'feedback effect function'. Depending on whether a specific behavioural hypothesis  $h_j$  is consistent with the observed data and whether

the participant's choice is correct in the current trial, the effect of the choice result on a specific latent hypothesis  $h_i'$  would be different. Therefore, I designed this probability function by considering four different conditions as detailed below.

(1) If the corresponding stimulus of  $h_j$  is in the chosen box, and the participant's choice is correct, then the participant would think that the corresponding latent hypothesis  $h_j'$  (note the subscript  $j$ ) is very likely to be related to the correct response (Figure 6.5). Therefore, the probability of the latent hypothesis  $h_j'$  should be much higher than those of the other latent hypotheses  $h_i'$ , i.e.,  $\alpha$  in the following equation (Equation 6.4) should be set to a high value close to 100%. Here I set  $\alpha = 0.98$ .

$$P(H' = h_i' | \{H = h_j, \mathbf{E}\}) = \begin{cases} \alpha & \text{if } h_i' \text{ corresponds to } h_j \\ (1 - \alpha)/(n + m - 1) & \text{otherwise} \end{cases} \quad (6.4)$$

In general  $n$  and  $m$  are the numbers of hypotheses which are respectively consistent and inconsistent with the choice of the current trial.  $n + m - 1$  appears in Equation (6.4) in order to make the sum of  $P(H' = h_i' | \{H = h_j, \mathbf{E}\})$  over all latent hypotheses  $h_i'$  equal to 1.0.

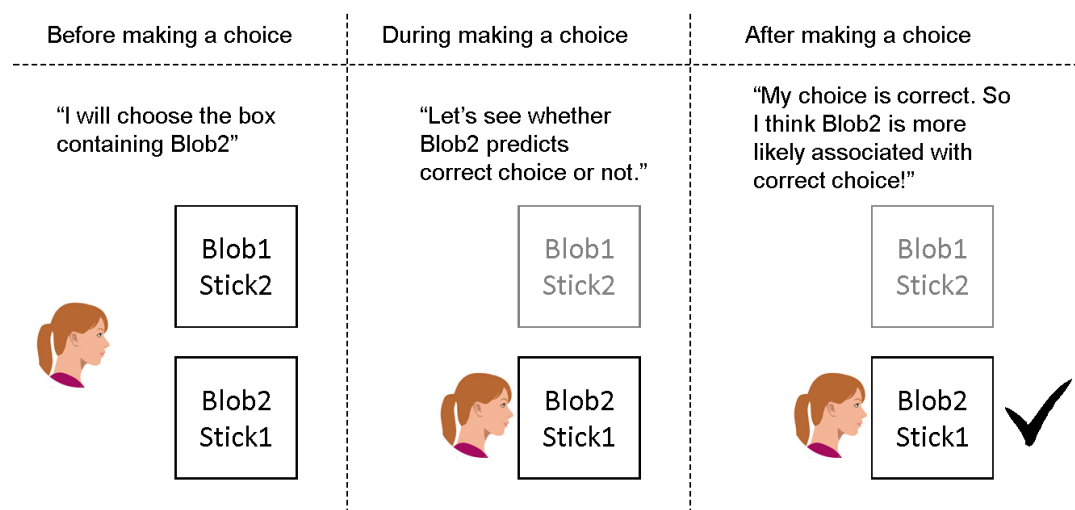


Figure 6.5: an example demonstrating the effect of feedback when a participant made a correct choice.

(2) If the corresponding stimulus of  $h_j$  is in the chosen box, and the participant's

choice is incorrect, then the participant can exclude both stimuli as being correct. In this case, the participant would think that every latent hypothesis  $h'_i$  consistent with the choice is not likely to be related to the correct response, and the correct response is more likely to be related to any latent hypothesis consistent with the stimuli in the un-chosen box (Figure 6.6); the probability function could, therefore, be designed as

$$P(H' = h'_i | \{H = h_j, E\}) = \begin{cases} (1 - \alpha) / n & \text{if } h'_i \text{ consistent with observation } E \\ \alpha / m & \text{if } h'_i \text{ inconsistent with observation } E \end{cases} \quad (6.5)$$

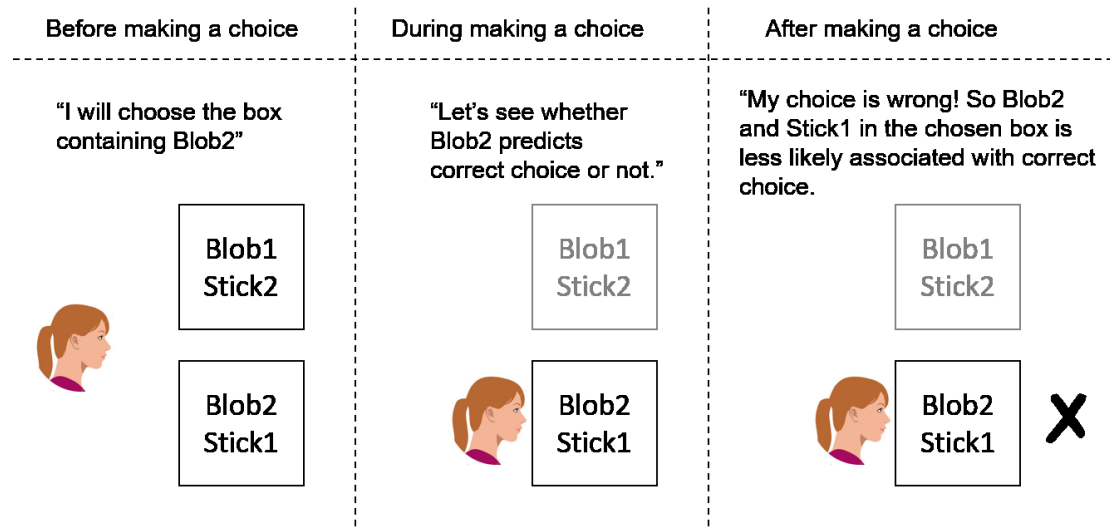


Figure 6.6: an example demonstrating the effect of feedback when a participant made an incorrect choice.

(3) If the corresponding stimulus of  $h_j$  is not in the chosen box, and the participant's choice is correct: in this case, any latent hypothesis  $h'_i$  consistent with the choice could be related to the correct response, while any other associative hypothesis  $h'_i$  inconsistent with the choice would be much less likely to be related to the correct response (Figure 6.7). So the function could be designed as

$$P(H' = h'_i | \{H = h_j, E\}) = \begin{cases} \alpha / n & \text{if } h'_i \text{ consistent with observation } E \\ (1 - \alpha) / m & \text{if } h'_i \text{ inconsistent with observation } E \end{cases} \quad (6.6)$$

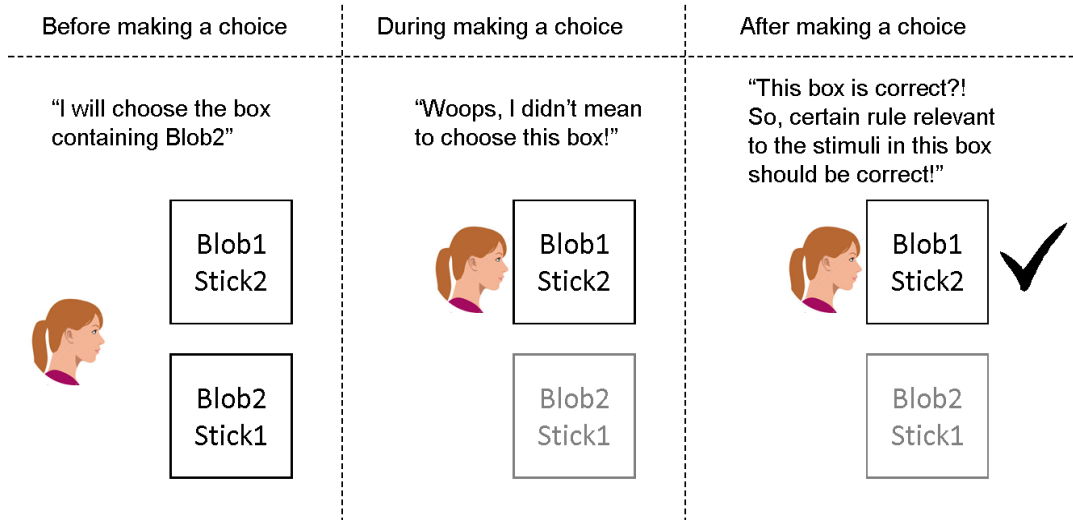


Figure 6.7: An example demonstrating the effect of feedback when a participant chose the correct box although she planned to choose the other box.

(4) If the corresponding stimulus of  $h_j$  is not in the chosen box, and the participant's choice is incorrect (Figure 6.8): this condition is similar to the condition (2) above, so

$$P(H' = h'_i | \{H = h_j, \mathbf{E}\}) = \begin{cases} (1 - \alpha) / n & \text{if } h'_i \text{ consistent with observation } E \\ \alpha / m & \text{if } h'_i \text{ inconsistent with observation } E \end{cases} \quad (6.7)$$

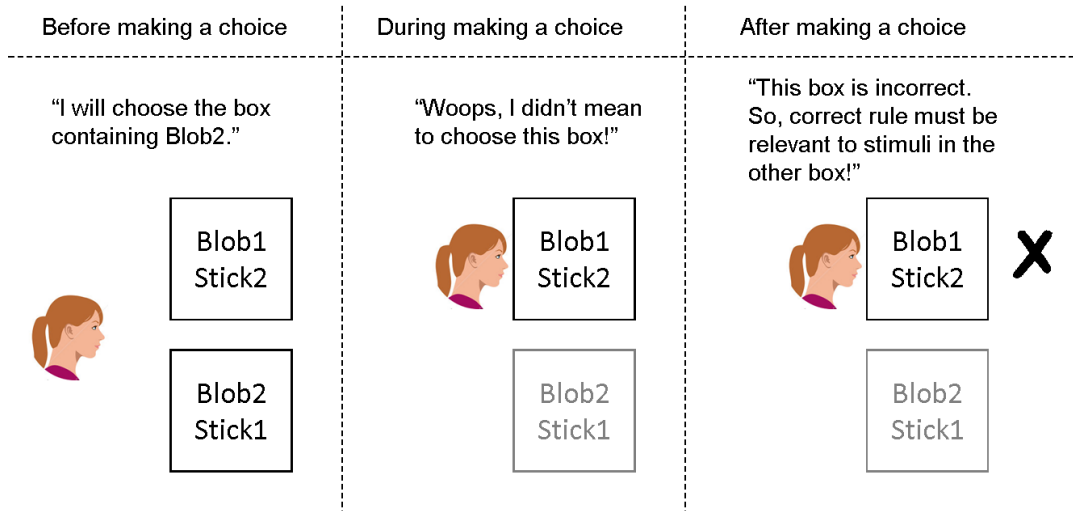


Figure 6.8: An example demonstrating the effect of choice result when a participant chose the incorrect box although she planned to choose the other box.

## Human behavioural reward Bayesian model $P(H = h_j | \mathbf{E})$

$P(H = h_j | \mathbf{E})$  measures how likely it is that the participant is using the specific behavioural hypothesis  $h_j$  to make their choice. Like the human behavioural simple Bayesian model above (Equation 6.1),  $P(H = h_j | \mathbf{E})$  can be calculated by

$$P(H = h_j | \mathbf{E}) = \frac{P(E | \{H = h_j, \mathbf{E}^*\}) \cdot P(H = h_j | \mathbf{E}^*)}{P(E | \mathbf{E}^*)} \quad (6.8)$$

The only difference between this Bayesian model and the human behavioural simple Bayesian model is in the design of the prior probability  $P(H = h_j | \mathbf{E}^*)$ . Intuitively, before making their choice at the current trial, when the participant thinks a specific latent hypothesis  $h_j'$  is more likely to be related to the correct response than any other latent hypotheses, the participant would be more likely to use the corresponding response pattern linked to the behavioural hypothesis  $h_j$  (note subscript is the same between  $h_j$  and  $h_j'$ ) to make their choice rather than use any other behavioural hypotheses. So here I use the posterior probability of the *latent hypothesis* at the previous trial to represent the prior probability of the corresponding behavioural hypothesis at the current trial. As a result, the prior probability has implicitly considered the effect of feedback (or reward) from previous trials. This is different to the design of the prior function in the rat behavioural Bayesian model (Section 2) and the human behavioural simple Bayesian model, where the posterior probability of the *behavioural hypothesis* in the previous trial is used to represent the prior probability of the same behavioural hypothesis in the current trial. In view of this, I call Equation (6.8) the *human behavioural reward Bayesian model*. See Table 6.3 for more details about these different rat and human models.



Table 6.3: development of Bayesian analyses

Bayesian analyses with parameters	Hypotheses	Properties	Dataset
<b>Rat behavioural simple Bayesian model:</b> $P(H = h_i   \mathbf{E}) = \frac{P(E   \{H = h_i, \mathbf{E}^*\}) \cdot P(H = h_i   \mathbf{E}^*)}{P(E   \mathbf{E}^*)}$ likelihood $P(E   \{H = h_i, \mathbf{E}^*\}) = \begin{cases} \gamma, & \text{if } h_i \sim E \\ 1 - \gamma, & \text{otherwise} \end{cases}; \gamma = 0.85$ prior $P(H = h_i   \mathbf{E})$ is from previous trial's posterior; undergraduate work	9 behavioural hypotheses: left bias, right bias, spontaneous alternation, win-stay, win-shift, M1 (Medium 1), M2, O1 (odour 1), O2	'1 <sup>st</sup> /2 <sup>nd</sup> ' observation was not considered in likelihood function; reward effect not considered; model for rats' behavioural patterns.	16 rats' 4ID data
<b>Rat behavioural Bayesian model:</b> $P(H = h_i   \mathbf{E}) = \frac{P(E   \{H = h_i, \mathbf{E}^*\}) \cdot P(H = h_i   \mathbf{E}^*)}{P(E   \mathbf{E}^*)}$ <div style="display: flex; justify-content: space-between;"> <div style="width: 48%;">                     Likelihood function for spatial hypotheses:  <math>P(E_2, E_1 = 1^{st}   \{H = h_i, \mathbf{E}^*\}) = \begin{cases} \alpha_1, &amp; \text{if } E_2 \sim h_i \\ \alpha - \alpha_1, &amp; \text{otherwise} \end{cases}</math>  <math>P(E_2, E_1 = 2^{nd}   \{H = h_i, \mathbf{E}^*\}) = \begin{cases} \alpha_2, &amp; \text{if } E_2 \sim h_i \\ 1 - \alpha - \alpha_2, &amp; \text{otherwise} \end{cases}</math>  <math>\alpha = 0.9 \quad \alpha_1 = 0.8 \quad \text{and} \quad \alpha_2 = 0.05</math> </div> <div style="width: 48%;">                     Likelihood function for perceptual hypotheses:  <math>P(E_2, E_1 = 1^{st}   \{H = h_i, \mathbf{E}^*\}) = \begin{cases} \beta_1, &amp; \text{if } E_2 \sim h_i \\ 0.5 - \beta_1, &amp; \text{otherwise} \end{cases}</math>  <math>P(E_2, E_1 = 2^{nd}   \{H = h_i, \mathbf{E}^*\}) = \begin{cases} \beta_2, &amp; \text{if } E_2 \sim h_i \\ 0.5 - \beta_2, &amp; \text{otherwise} \end{cases}</math>  <math>\beta_1 = \beta_2 = 0.45</math> </div> </div>	8 behavioural hypotheses: spontaneous alternation, perseveration, win-stay, win-shift, M1, M2, O1, O2.	'1 <sup>st</sup> /2 <sup>nd</sup> ' observation was considered in likelihood function; impact of feedback not considered; model only for rats' behavioural patterns.	93 rats' 7-stage data
<b>Human behavioural simple Bayesian model:</b> the same as the first row above	10 behavioural hypotheses: blob 1, blob 2, stick 1, stick 2, blob1 and blob 2 alternate, stick 1 and stick 2 alternate, blob 1 + stick 1, blob 1 + stick 2, blob 2 + stick 1, blob 2 + stick 2	'1 <sup>st</sup> /2 <sup>nd</sup> ' observation not considered; reward effect not considered; model for humans' behavioural patterns.	32 humans' 7-stage data
<b>Human latent probabilistic model:</b> $P(H' = h'_i   \mathbf{E}) = \sum_{h_j} P(H' = h'_i   \{H = h_j, \mathbf{E}_i\}) \cdot P(H = h_j   \mathbf{E})$ $P(H = h_j   \mathbf{E})$ is a <b>human behavioural reward Bayesian model</b> , wherein prior function is from previous	10 associative hypotheses: blob 1, blob 2, stick 1, stick 2, blob1 and blob 2 alternate, stick 1 and stick 2 alternate, blob 1 + stick 1; blob 1 + stick 2; blob 2 +	'1 <sup>st</sup> /2 <sup>nd</sup> ' not considered; reward effect considered; model for both humans' behaviour and reward association; conceptually,	32 humans' 7-stage data

trial's $P(H'=h'_i   \mathbf{E})$ which considers effect of feedback/reward.	stick 1; blob 2 + stick 2	also applicable to rats.	
<p>Feedback effect function <math>P(H'=h'_i   H=h_j, \mathbf{E})</math>: <math>\alpha = 0.98</math>; <math>n=3</math>, <math>m=5</math></p> <p>If <math>h_j \sim E</math> &amp; choice is correct: <math>P(H'=h'_i   \{H=h_j, \mathbf{E}\}) = \begin{cases} \alpha &amp; \text{if } h'_i \propto h_j \\ (1-\alpha)/(n+m-1) &amp; \text{otherwise} \end{cases}</math></p> <p>If not <math>h_j \sim E</math>, &amp; choice is correct: <math>P(H'=h'_i   \{H=h_j, \mathbf{E}\}) = \begin{cases} \alpha/n &amp; \text{if } h'_i \sim E \\ (1-\alpha)/m &amp; \text{otherwise} \end{cases}</math></p> <p>If choice is incorrect: <math>P(H'=h'_i   \{H=h_j, \mathbf{E}\}) = \begin{cases} (1-\alpha)/n &amp; \text{if } h'_i \sim E \\ \alpha/m &amp; \text{otherwise} \end{cases}</math></p>	<p>10 behavioural hypotheses:  blob 1, blob 2, stick 1, stick 2,  blob1 and blob 2 alternate, stick 1  and stick 2 alternate, blob 1 +  stick 1, blob 1 + stick 2, blob 2 +  stick 1, blob 2 + stick 2</p>		

## 6.5 Data screening

Thirty-eight participants participated and finished the human 7-stage task. While most participants finished the task within 100 trials, four participants took more than 200 trials which made them appear to be outliers ( $z > 2$ ; Figure 6.9, left). Therefore, I did not analyse the four cases on the following data analysis. For the remaining 34 cases, I further collected the maximum stage trials from each participant's data and found that one participant took more than 90 trials to finish a single stage (more specifically the ID stage; see Figure 6.9, right), which, again, is much higher than the other participants' maximum stage trials ( $z > 3$ ). In addition, one participant wrote down everything observed in each trial during learning; this participant's data was removed as well, therefore. As a result, this participant's data were also excluded, resulting in a final sample size of  $n=32$  participants.

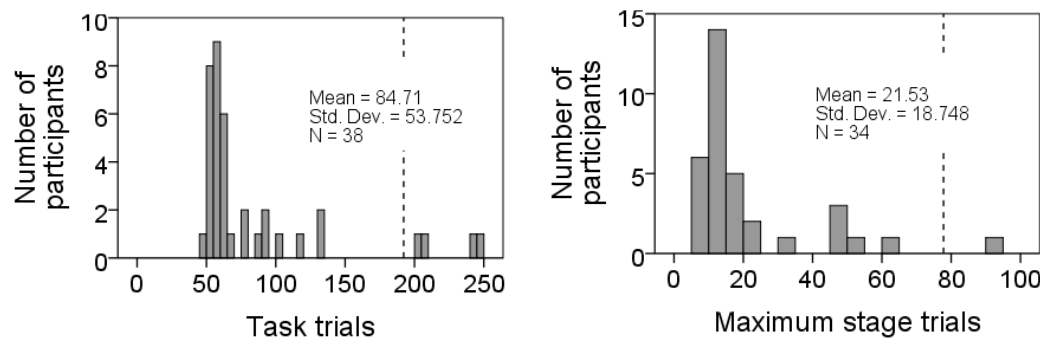


Figure 6.9: the histogram of task trials over all 38 data (left) and the histogram of the maximum stage trials after removing four outliers (right). The data on the right side of the dashed vertical lines ( $z=2$  in the left figure and  $z=3$  in the right figure) were considered outliers and were removed.

In addition, when analysing participants' confidence, I excluded the data of those participants whose confidence was likely not to have been properly reported. In the remaining 32 participants' data, visual inspection of the reported confidence found that two participants reported confidence '0' (i.e., the minimum possible value) in almost every trial in each stage (even in the last six correct trials), suggesting that these two participants did not respond to the question about confidence in the entire task. The other two participants reported confidence '100' (i.e., the maximum possible value) in almost every trial in each stage, suggesting that the two participants probably failed to understand the instruction on answering the confidence question. As a result, 28 participants' confidence data were remained after excluding the four participants' data, as described above. Note that such confidence data exclusion is applied only to the analysis about participants' confidence (Section 7.4).

## 7 Results

### 7.1 Trials to criterion

One-way between-subject ANOVA showed that there was a significant effect of learning stage on the trials to criterion,  $F(6, 224)=8.32$ ,  $p<0.001$ , partial  $\eta^2=0.182$ . Pairwise post-hoc  $t$ -tests with Dunn-Sidák's correction indicated that trials to criterion in the ED stage were significantly higher than those in all the other six stages, while there is no significant difference in the trials to criterion between any two of the other six stages (Figure 7.1), as though the participants formed an attentional set before the ED stage. Therefore, similar to findings in the rat 7-stage task, humans take significantly more trials to learn the ED stage compared to the ID stage. Note that although participants often took more trials to criterion in REV1 than in the CD, the difference in trials between the two stages was not significant, probably because human participants could quickly change response patterns to make their choice when spotting wrong choices early in the REV1 stage.

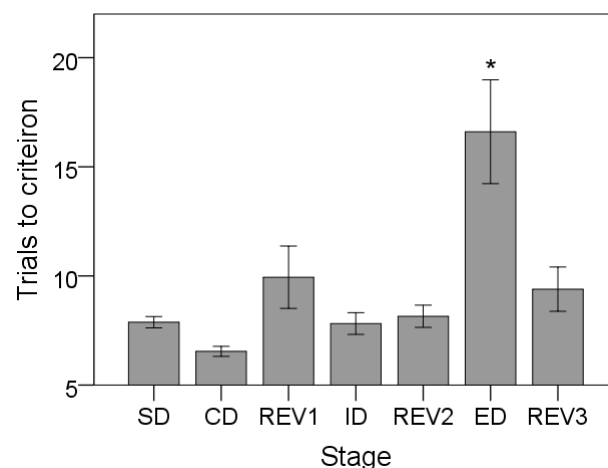


Figure 7.1: trials to criterion for each of the seven stages. Humans took significantly more trials in ED than that in ID stage.

### 7.2 Can Bayesian model accurately estimate which erroneous rules were used?

One main purpose of designing the human 7-stage task is to check whether a participant really did use a specific response pattern to make their choice when the Bayesian estimate of the corresponding hypothesis was high (see Section 2.3 for the relationship between response pattern and corresponding hypothesis). Since both the human behavioural simple Bayesian model (Equations 6.1-6.2) and the human behavioural reward Bayesian model (Equation 6.8) can be used to estimate what response patterns (sometimes called *rules* in the research literature) have been used to make choices, these two Bayesian models were evaluated as to how reliably they estimated response patterns, particularly the reward-irrelevant (i.e.,

erroneous) response patterns, in this section. Specifically, if a Bayesian model estimated that one irrelevant response pattern was used (i.e., the Bayesian probability of the corresponding hypothesis is larger than 0.6) in *one or more trials* in a learning stage, and the participant also did use this response pattern (known from participant's oral response) in the stage, then we would consider that the Bayesian model has *correctly* found (or estimated) *one* erroneous response pattern; otherwise, the Bayesian model has *incorrectly* estimated *one* erroneous response pattern. Note that even if the probability of one type of response pattern is larger than 0.6 in more than one trial in a stage, we only count once for this type of response pattern in the stage. The focus of data analysis here is to compare which Bayesian model can correctly estimate more erroneous response patterns. Across all participants' seven learning stages, the human behavioural simple Bayesian model correctly estimated 155 irrelevant response patterns and incorrectly estimated 21 irrelevant response patterns (Table 7.1), after removing four uncertain estimates due to participants' ambiguous oral responses, such as 'I guess this one is correct' or 'I want to see what happens'. In comparison, the human behavioural reward Bayesian model correctly estimated 181 hypotheses and incorrectly estimated just seven hypotheses (Table 7.1, last row), after removing three uncertain estimates. A chi-square test supports that the human behavioural reward Bayesian model found relatively more correct estimates of response patterns and fewer incorrect estimates than the human behavioural simple Bayesian model, Chi-square (1) = 8.63,  $p=0.003$ . Note that the total number of estimated erroneous response patterns is different between the two Bayesian models (176 vs 188), because the human behavioural simple Bayesian models missed more erroneous rules used by rats (while it found more false erroneous rules).

Table 7.1: number of reward-irrelevant response patterns correctly and incorrectly estimated by two Bayesian models, across all participants' seven learning stages, with percentage of correct and incorrect estimates in parentheses and expected values in italics.

	Correct estimates of irrelevant response patterns	Incorrect estimates of irrelevant response patterns
Human behavioural simple Bayesian model	155 (88.1%) <i>162.5</i>	21 (11.9%) <i>13.5</i>
Human behavioural reward Bayesian model	181 (96.3%) <i>173.5</i>	7 (3.7%) <i>14.5</i>

A mixed design two-way ANOVA analysis showed that the number of correctly estimated irrelevant (i.e., erroneous) response patterns, when using the human behavioural reward Bayesian model, is significantly higher than using the simple Bayesian model,

$F(1,224)=17.56$ ,  $p<0.001$ , partial  $\eta^2 = 0.073$ . Also, the number of correctly estimated erroneous response patterns is significantly different between the seven stages,  $F(6,224)=48.91$ ,  $p<0.001$ , partial  $\eta^2 = 0.567$ . The interaction of model with stage is significant,  $F(6,224) = 15.74$ ,  $p<0.001$ , partial  $\eta^2= 0.297$ .

Post-hoc Tukey's HSD tests showed that significantly more irrelevant response patterns are found in the ED and three reversal stages than in the other three stages (SD, CD, ID), and significantly fewer irrelevant hypotheses are found in the CD and ID stages than in the SD stage, all at the .05 level of significance. All other comparisons were not significant (Figure 7.2).

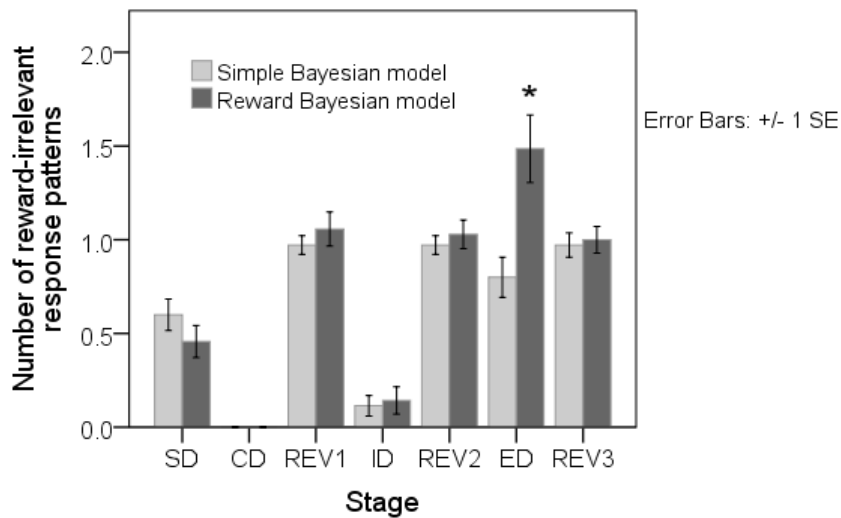


Figure 7.2: the number of reward-irrelevant (i.e., erroneous) response patterns correctly estimated by the two Bayesian models for each learning stage. The human behavioural reward Bayesian model can find more irrelevant (erroneous) hypotheses used by participants in the ED stage.

The human behavioural reward Bayesian model found significantly more irrelevant hypotheses in the ED stage than the simple Bayesian model does (Figure 7.2), paired  $t(32) = 5.81$ ,  $p<0.001$ ,  $d=1.03$ . For example, the reward Bayesian model has correctly estimated three erroneous response patterns in the ED stage (pointed by arrows in Figure 7.3, left) from the data of participant 'A003': two response patterns from the irrelevant dimension (solid and dotted black curves) and the stick alternation hypothesis (solid red curve), while the simple Bayesian model only estimated one erroneous response pattern from the irrelevant dimension (dotted black curve pointed by arrow; Figure 7.3, right).

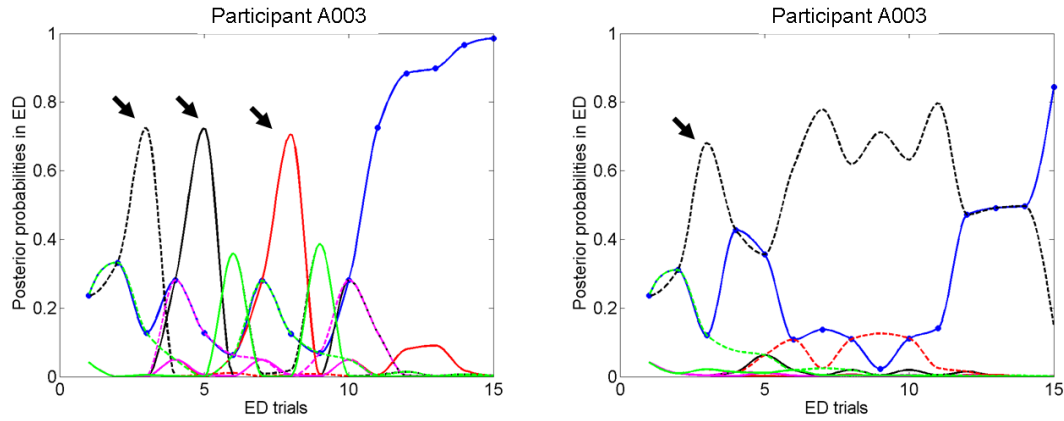


Figure 7.3: an illustrative example of the behavioural pattern estimated by the human behavioural reward Bayesian model (left) and by the human behavioural simple Bayesian model (right) on the same data from participant ‘A003’. The reward Bayesian model has correctly estimated more erroneous hypotheses (left), where the three estimated hypotheses were found to be used in this participant’s oral report. Solid and dotted black curves: probabilities of two hypotheses from the previously rewarded but currently irrelevant dimension; Solid blue curve: rewarded stimulus; Dotted blue curve: non-rewarded stimulus from the currently relevant dimension; Solid and dotted red curves: stick alternation and blob alternation hypotheses; Solid and dotted magenta and green curves: four blob-stick combinations.

### 7.3 Individual estimate of set-formation before the ED stage

Bayesian analysis provides a way to check whether *individual* participants have formed an attentional set. The rationale is that if a participant has formed an attentional set before starting the ED stage this participant would use (new) stimuli of the previously relevant but currently irrelevant dimension to make choices at the first several (e.g., 4) trials in the ED stage. This would lead to a relatively high Bayesian estimate for such irrelevant stimuli at the beginning of ED stage, if the human behavioural reward Bayesian model works well.

The human behavioural reward Bayesian model (Equation 6.8) can accurately estimate that participants had formed an attentional set when they really did. For 27 out of 32 participants’ data, the posterior probability of one or two hypotheses from the previously rewarded but currently irrelevant dimension becomes relatively high (i.e.,  $>0.6$ ) at the first several trials in the ED stage (see Figure 7.4 for two representative examples). Based on the participants’ oral response to the question ‘why do you think this is correct?’, it appears that these 27 participants *did use* stimuli from the previously relevant dimension to make choices at the beginning of the ED stage, indicating that these participants had formed an attentional set before starting the ED stage.

What is more, for the remaining five participants for whom the posterior probability of each stimulus from the previously reward-relevant dimension is relatively low (all less than 0.4) at the first several trials in ED (Figure 7.5), these participants' oral responses tell us that they *really did not* make choices based on the previously reward-relevant dimension at the beginning of the ED stage. More specifically, one participant 'A012' used 'how much the stick is outside the blob region' to make choices (Figure 7.5, left), whereas the other 4 participants immediately used the currently reward-relevant dimension to make choices at the beginning of the ED stage (Figure 7.5, right).

The perfect match between Bayesian estimate and participants' oral responses suggests that the reward Bayesian model can be used to analyse whether each individual participant has formed an attentional set or not from the first several trials in the ED stage.

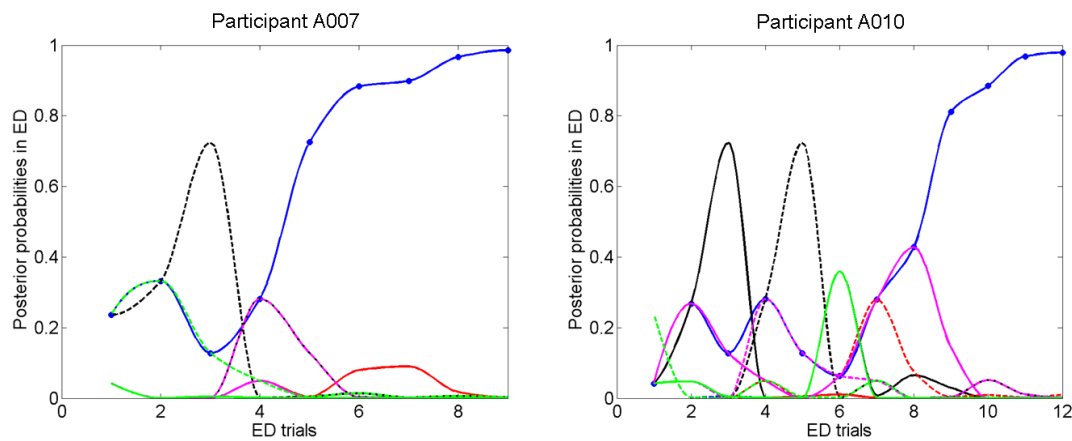


Figure 7.4: posterior probability of each hypothesis over ED trials by the human behavioural reward Bayesian model from two representative participants who made choices based on the previously relevant dimension's stimuli at the beginning of the ED stage. Solid and dotted black curves: two stimuli from the previously relevant but currently irrelevant dimension; Solid blue curve: rewarded stimulus; Dotted blue curve: non-rewarded stimulus from the currently relevant dimension; Solid and dotted red curves: stick alternation and blob alternation hypotheses; Solid and dotted magenta and green curves: four blob-stick combinations.



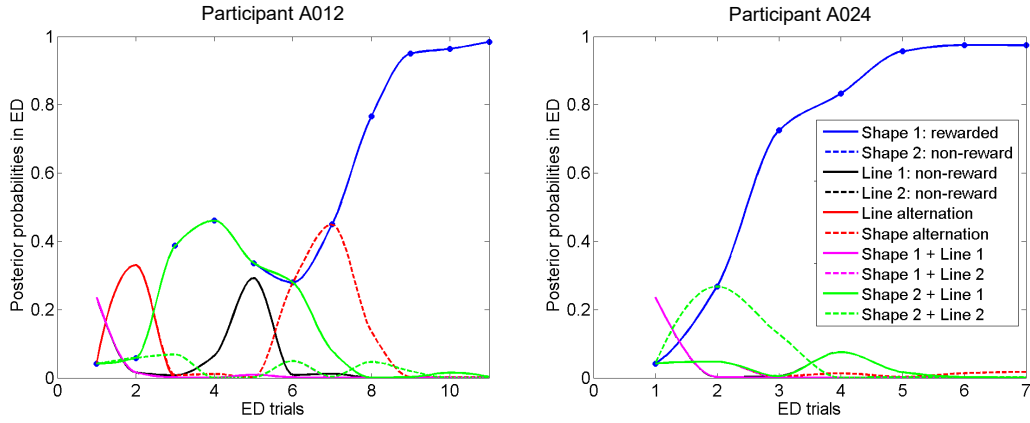


Figure 7.5: posterior probability of each hypothesis over ED trials by the human behavioural reward Bayesian model from two participants who did not make choices based on stimuli (the sticks in the stimulus) in the previously relevant dimension at the beginning of the ED stage. The Bayesian estimates of the two hypotheses from the previously relevant dimension (two black sticks) were less than 0.6 over all trials in each ED stage.

In comparison, with the human behavioural simple Bayesian model, for those 27 participants who made choices based on the newly irrelevant dimension's stimuli at the beginning of the ED stage, the posterior probabilities of certain hypotheses from the irrelevant dimension become relatively high on only 16 participants' data. In other words, the simple Bayesian model failed to find that 11 participants focused on the previously relevant but currently irrelevant dimension at the beginning of the ED stage. For those five participants who did not make their choices based on the irrelevant dimension's stimuli at the beginning of the ED stage, the posterior probability of a hypothesis from the irrelevant dimension becomes relatively high on one participant's data. This result suggests that the simple Bayesian model, which did not consider the effect of choice feedback, cannot reliably estimate whether or not each participant has formed an attentional set.

#### 7.4 Latent probabilistic model can more accurately estimate participants' confidence

The human latent probabilistic model (Equation 6.3) was developed to estimate the likelihood that each hypothesis is associated with reward (or 'correct') by considering the effect of the choice result ('correct' or 'incorrect') on consequent learning. During learning to find the 'correct'-associated hypothesis, participants' confidence in making choices would probably be affected by their choice feedback ('correct' or 'incorrect') from previous trials, which is supported by the result shown in Figure 7.6. Therefore, compared to the human behavioural simple Bayesian model, which can only estimate the likelihood each hypothesis is used to make choice, the reward Bayesian model may more reliably estimate participants' confidence

in making choices.

Since participants took, on average, just 6 to 10 trials to finish most learning stages except for the ED stage, the pattern of participants' confidence in these six stages may not be rich enough to show potential differences in confidence estimation between the behavioural simple Bayesian model and the latent probabilistic model. Therefore, I analysed only the confidence data in the ED stage. With the latent probabilistic model, the correlation between the reported confidence and the posterior probability (i.e., b-value) of the correct (latent) hypothesis in the ED stage was calculated for each of the 21 participants, resulting in 21 confidence-b-value correlations. Similarly, with the simple Bayesian model, the correlation between the reported confidence and the probability estimate of the correct (behavioural) hypothesis was calculated for each participant. A Wilcoxon signed-rank test showed that the confidence-b-value correlation from the latent probabilistic model is significantly higher than that from the simple Bayesian model ( $Z=2.41$ ,  $p=0.016$ ,  $r=0.32$ ), median=0.74 with latent probabilistic model and median=0.62 with simple Bayesian model (Figure 7.7). Figure 7.8 shows an example of the scatter plot of b-value versus reported confidence with the latent probabilistic model and the simple Bayesian model, respectively.

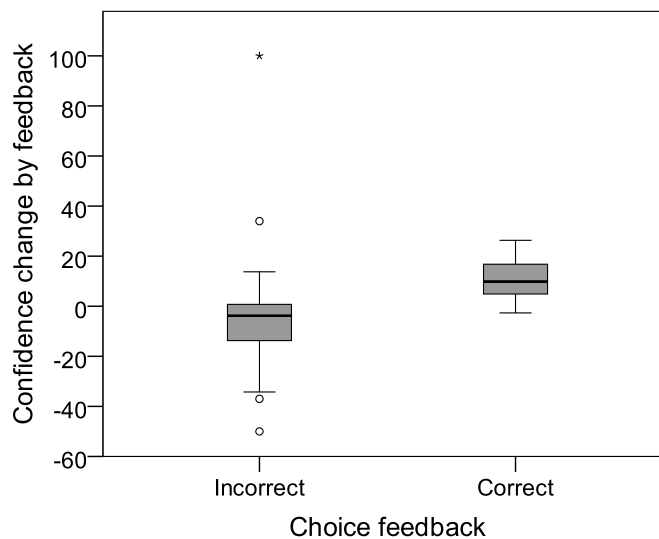


Figure 7.6: effect of choice feedback in current trial on the change in participants' choice confidence in the next trial. The average changes in participants' reported confidence (within the range 0 to 100) for both 'correct' and 'incorrect' feedbacks were collected over trials in the ED stage for each participant. A Mann-Whitney U test showed that the change in participants' choice confidence is significantly different between the 'correct choice' feedback and the 'incorrect choice' feedback (median change in confidence is 9.8 when choice is 'correct', and -3.8 when choice is 'incorrect'),  $U=121.0$ ,  $n_1=28$ ,  $n_2=28$ ,  $Z=-4.44$ ,  $p<0.001$ .

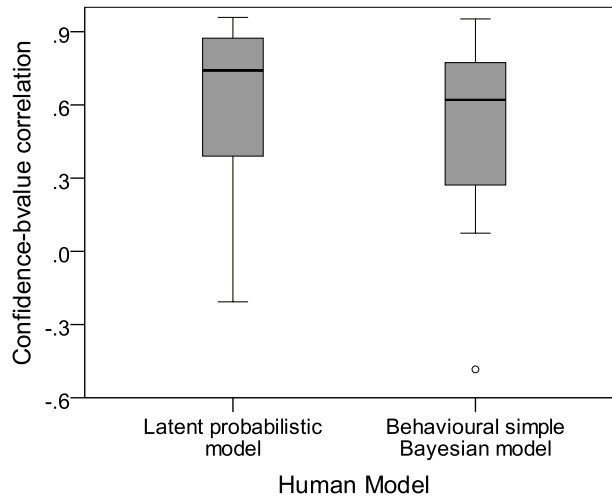


Figure 7.7: the correlation between reported confidence and the b-value with the human latent probabilistic model and the simple Bayesian model, respectively. The b-value from the latent probabilistic model is more correlated with the reported confidence than that from the simple Bayesian model.

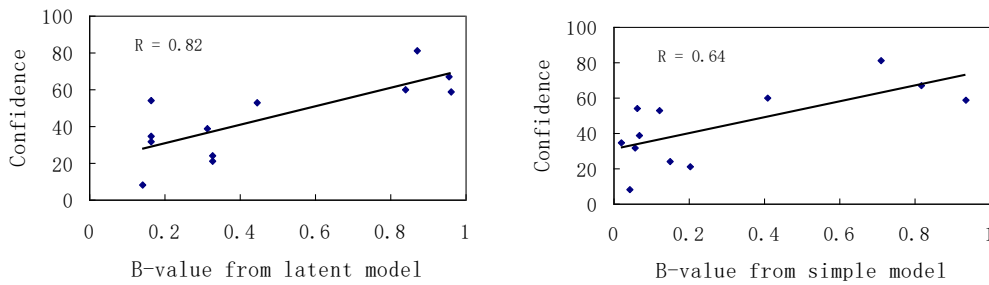


Figure 7.8: an illustrative example to show that the b-value from the human latent probabilistic model (left) is more strongly correlated with the reported confidence than that from the simple Bayesian model (right).

## 7.5 Bayesian learning criterion based on the latent probabilistic model performs better than the simple Bayesian model

As discussed for the rat Bayesian model, a Bayesian learning criterion can be used as an alternative learning criterion to decide the point at which participants have learned the correct rules. Since we know what actual hypotheses participants are using for each learning trial, based on their oral report for each trial, we can validate how well the Bayesian learning criterion is based both on the human behavioural simple Bayesian model and on the latent probabilistic model, and also validate how well the 6-in-a-row criterion is.

From the participants' oral reports, I found that 29 out of 32 participants started to use the correct rules from either the sixth last, fifth last, fourth last, or the third last trials in each of the seven learning stages. This suggests that the 6-in-a-row criterion is well aligned with the point at which participants learned the correct rules.

For the other three participants, a total of four false positives were found from their ED or reversal stages, whereas participants in general made the last six correct choices based on the combination or interaction of blob and stick stimuli, like “more proportion of sticks insides the blob region”.

Based on the human behavioural simple Bayesian model, I found that the Bayesian learning criterion was not satisfied in the last trial of the three reversal stages for most participants and in the last trial of the ED stage for about half of the participants (Figure 7.9). This suggests that the Bayesian learning based on the simple model is not consistent with participants’ oral reporting of the point at which they have learned the correct response patterns.

In comparison, with the human latent probabilistic model, Bayesian learning criterion is satisfied in the last trial of all stages for all participants (Figure 7.10). Detailed analysis also shows that the criterion is satisfied within the last three trials for ID, ED, and reversal stages. This suggests that the Bayesian learning criterion based on the latent probabilistic model is well aligned with the 6-in-a-row criterion and consistent with participants’ oral report of when they have learned the rules for all learning stages except the six false positive learning stages.

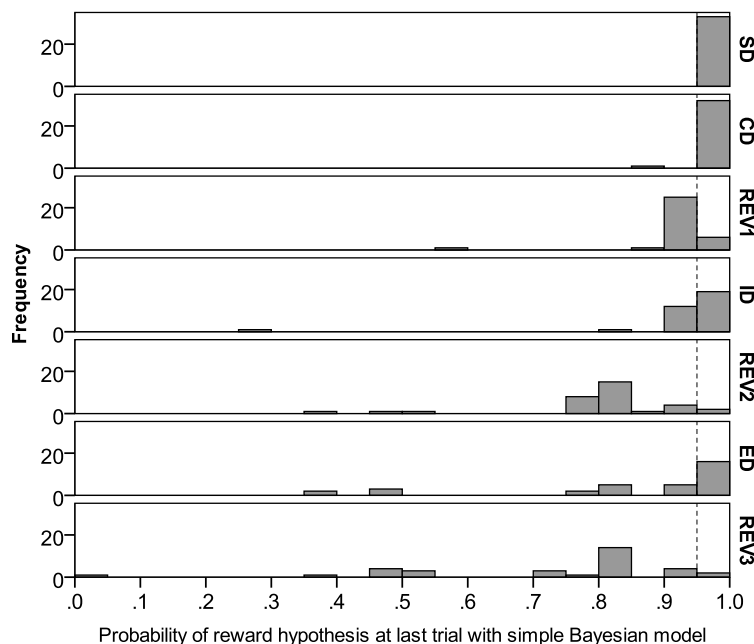


Figure 7.9: histogram of probability of ‘correct’-associated hypotheses in the last trial for each learning stage over 32 participants with the human behavioural simple Bayesian model. The Bayesian learning criterion based on the simple Bayesian model is not working well in at least the three reversal stages and the ED stage.

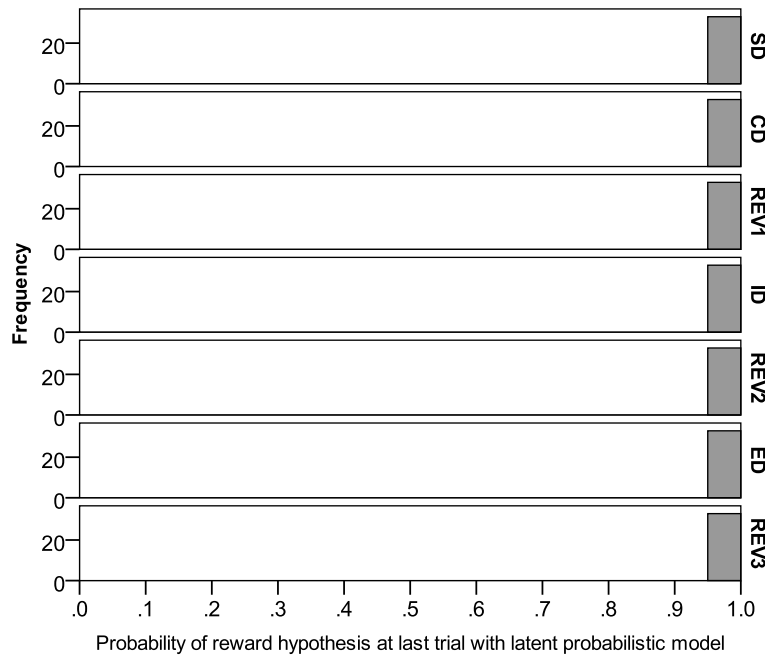


Figure 7.10: histogram of probability of ‘correct’ latent hypotheses at the last trial for each learning stage over 33 participants with the human latent probabilistic model. The Bayesian learning criterion based on the latent probabilistic model is well aligned with participants’ oral reports of the point at which they learned the correct rules.

## 7.6 Human latent probabilistic model better predicts whether participants are using correct response patterns to make choices

After showing that the human latent probabilistic model can better predict the point at which participants started to use the correct response pattern to make their choices, we investigated whether the b-value (i.e., the posterior probability of the correct latent hypothesis) from the human latent probabilistic model performs better than the b-value from the human behavioural simple Bayesian model in predicting whether a participant is using the correct response pattern to make their choice in each (current) trial in the ED stage. Specifically, the b-value and ‘whether or not participants are using the correct response pattern to make their choice’ (obtained from participants’ oral reports) were collected for each trial in the ED across all 32 participants. Then a logistic regression was fitted to predict ‘whether or not participants are using the correct response patterns to make their choice’ with b-value as the predictor. Similarly, another logistic regression was fitted using the b-value of the simple Bayesian model as the predictor.

For both predictors, the residual of prediction for each trial was computed as the difference between the predicted value (within the range 0 to 1) and the actual value (with ‘1’ for ‘using correct response pattern to make their choice’ and ‘0’ otherwise). If the b-value from the

human behavioural simple Bayesian model and the b-value from the human latent probabilistic model have similar prediction performance, the mean absolute residuals over all trials per participant from both predictors should be similar when predicting the actual value. A Wilcoxon signed-rank test showed that the absolute residual from the human latent probabilistic model (median=0.20) is significantly lower than that from the human behavioural simple Bayesian model (median=0.28),  $Z=-4.77$ ,  $n=32$ ,  $p<0.001$ ,  $r=0.60$ . (Figure 7.11). This supports the view that the latent probabilistic model is better than the simple Bayesian model in predicting whether or not participants are using the correct response pattern to make their choice in each trial.

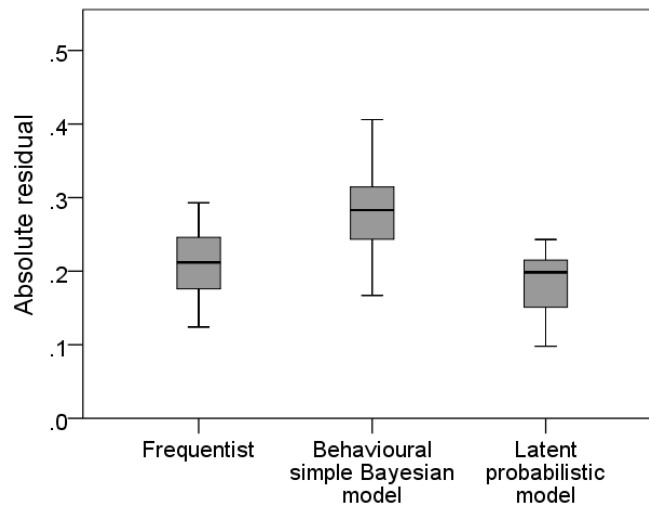


Figure 7.11: average absolute value of the residual per participant based on the frequentist p-value, b-value from the human behavioural simple Bayesian model, and b-value from the human latent probabilistic model.

In addition, I also fitted a logistic regression with the frequentist p-value as predictor. A Wilcoxon signed-rank test showed that the absolute residual based on p-value (median=0.21) is also significantly lower than that from the simple Bayesian model (median=0.28),  $Z=-4.06$ ,  $n=32$ ,  $p<0.001$ ,  $r=0.51$ , but is significantly higher than that from the latent probabilistic model,  $Z=3.33$ ,  $p<0.01$ ,  $r=0.42$  (Figure 7.11). These results suggest that although both the 6-in-a-row and the Bayesian learning criterion based on the human latent probabilistic model are aligned with the point at which participants started to use the correct response pattern to make their choice, the latent probabilistic model can better predict whether participants are using the correct response pattern to make their choice for each learning trial (including the trials before the last six correct trials).

## 7.7 The human behavioural reward Bayesian model does not work well on rats

Considering the good performance of the human behavioural reward Bayesian model on the

human task, it is interesting to explore whether the human reward Bayesian model performs well on the archival rat data used in the previous chapter (see Section 2.1). The evaluation of the human behavioural reward Bayesian model requires that we know what response patterns have been actually used by rats. However, we can never know for certain the rats' internal decision-making processes in any given learning stage. Even so, we may assume a rat used the *perseveration* response pattern when the rat directly chose bowls on the same side (either left side or right side) over six or more trials in a row, without going to check the other bowl. Note that in this case, the rat cannot make six correct choices in a row because the configurations of bowls were designed such that the reward-relevant stimulus would not appear on the same side for more than three trials in a row. Similarly, the rat probably won't make choice based on other perceptual response patterns. In this case, we may evaluate whether a Bayesian model can correctly estimate that the rat is actually using the perseveration response pattern to choose bowls. From all the 47 control rats, I observed that there are 28 stages where the rats chose bowls on the same side over six or more consecutive trials. For these 28 stages, the rat behavioural Bayesian model developed for the rat task (aka 'rat model') correctly estimated that the rat used the perseveration response pattern in all the 28 stages, while the human behavioural reward Bayesian model (aka 'human model') found the perseveration response pattern was used in only 11 stages (Table 7.2). Each Bayesian model estimated that a rat used the perseveration response pattern when the posterior probability of the corresponding hypothesis is larger than 0.6. A Chi-square test shows that the rat model performs significantly better than the human model in correctly estimating the perseveration hypothesis,  $\text{Chi-square}(1) = 24.4$ ,  $p < 0.001$ . Figure 7.12 shows an example where the rat model found that the rat used the perseveration response pattern (red dotted curve) while the human model did not, based on the criterion that the Bayesian estimate of the hypothesis is larger than 0.6. A similar result was obtained when the parameter in the human model was changed, precisely by varying the parameter in the feedback effect function (Equations 6.4 – 6.7). The results suggest that although the human model can accurately estimate the response patterns used by human participants in the human task it does not work well on the rat task. This is because the human model estimated that the rat would be more likely to use certain other response pattern(s) than the perseveration response pattern once the rat made a wrong choice, whereas the rat made choices on the same side even after making multiple wrong choices.

Table 7.2: estimate of perseveration response pattern (based on the threshold 0.6; see main text) with two Bayesian models from those stages where rats chose bowls from the same side on at least six consecutive trials.

	Correct estimate	Incorrect estimate
Rat behavioural Bayesian model (rat model)	28 (100%) (expected: 19.5)	0 (0%) (expected: 8.5)
Human behavioural reward Bayesian model (human model)	11 (39.3%) (expected: 19.5)	17 (60.7%) (expected: 8.5)

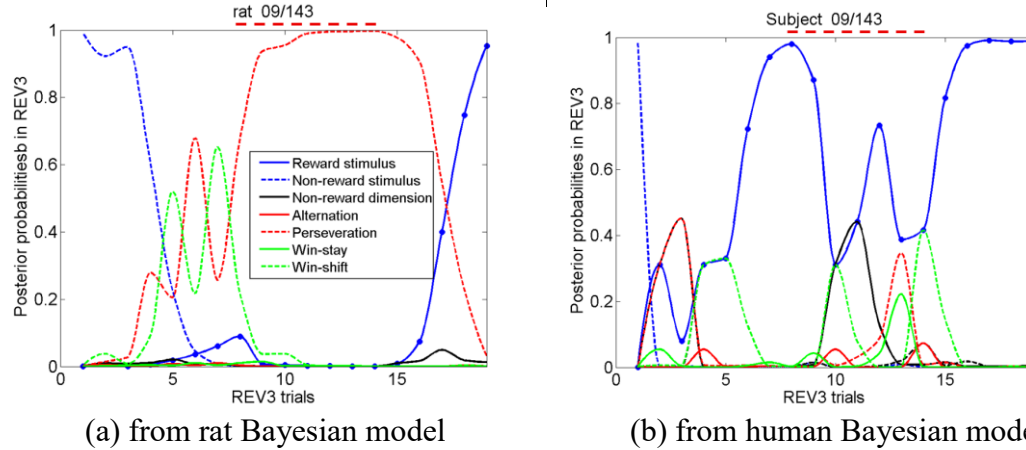


Figure 7.12: a representative example of the Bayesian estimate of the ‘perseveration hypothesis’ (red dotted curve) with the rat model (left) and the human model (right) when the rat (‘09/143’) probably used the spatial perseveration response pattern to make its choice over eight consecutive trials in the REV3 stage. The rat model correctly estimated the hypothesis, but the human model did not. A red dashed bar above each figure indicates trials in which rats used *perseveration* for bowl choice.

To further compare the rat model and the human model on rat data, I also evaluated their performance in predicting whether a rat’s next trial choice is correct or not based on the current trial’s b-value, where the b-value is either from the rat model or from the human model. Specifically, an individual logistic regression was fitted to predict the next trial’s choice correctness, based on the current trial’s b-value over all seven stages’ learning trials for *each* individual rat, resulting in 93 individual logistic regressions for each b-value predictor. The residual of prediction for each trial was computed as the difference between predicted value (within the range 0 to 1) and the actual correctness (with ‘1’ for correct choice and ‘0’ for incorrect choice).

If the b-value from the rat model and the b-value from the human reward Bayesian model have similar prediction performance, the mean absolute residuals over all trials per participant from both predictors should be similar when predicting the actual value. A paired t-test showed that the absolute residual from the human model ( $m=0.350$ ,  $sd=0.061$ ) is significantly lower than that from the rat model ( $m=0.357$ ,  $sd=0.060$ ),  $t(92)=-4.58$ ,  $p<0.001$ ,



$d=0.48$ . However, considering the very large absolute residual (0.350 vs. 0.357; the expected absolute residual from random guess is 0.500) and the very small difference ( $-0.007$ ) in average absolute residual between the two models, it suggests that the rat model and the human model have similar performance in predicting the next trial's outcome. A similar finding was obtained when the b-value is from the human latent probabilistic model, replacing the b-value from the human behavioural reward Bayesian model.

## **8 Discussion on Bayesian analysis of the human task**

The Bayesian analysis of the human task shows that the Bayesian estimates of both correct and erroneous response patterns match with what participants reported orally, providing indirect evidence for the validity of Bayesian analysis of rat data with a similar Bayesian model. For humans, the Bayesian analysis considering the effect of choice feedback is clearly better than the model without considering the effect in estimating participants' response patterns and choice confidence. However, the human reward Bayesian model does not offer any substantive advantage to analysing the archival rat data, in spite of doubling the free parameters used in the analysis relative to the rat Bayesian model in both cases. Thus, I believe that the enhanced predictive accuracy in humans of incorporating feedback into the analysis is not simply a statistical artefact. I speculate that humans and rats use different learning approaches, with humans using rapidly adapting hypothesis-testing and rats relying on more gradual associative learning.

### **8.1 Human reward Bayesian model can accurately estimate erroneous rules used by participants**

Participants may often try one or more types of erroneous, irrelevant response patterns before learning to find the correct, rewarded response patterns, particularly in the ED stage. If our Bayesian model(s) can find (at least some of) the erroneous response patterns used by participants, it would provide very useful information for investigating individuals' detailed learning processes. Both the human behavioural simple Bayesian model and the human behavioural reward Bayesian model have shown their ability to estimate the erroneous response patterns. Specifically, when the Bayesian estimate of one response pattern is high in an individual's learning stage, most of the time the estimate is correct by comparing with the participant's oral report. The estimate accuracy is 96.3% (181 correct estimates out of 188 response patterns) for the human behavioural reward model and 88.1% (155/176) for the human behavioural simple Bayesian model. The human behavioural reward model correctly

found more erroneous response patterns (particularly in the ED stage) with fewer mistakes than the human behavioural simple Bayesian model. When the human behavioural reward model made a mistake, it was either because the participant accidentally made a choice that he/she did not mean to, or because the participant used rare and special rules about the conjunction of blob and stick stimuli. This suggests potential future improvements in human task design, e.g., providing an opportunity for participants to change their choice at each learning trial, and carefully designing and pairing blob and stick stimuli such that the rare blob-stick interaction response patterns become invalid in differentiating the two choices at each trial.

In addition, the good estimate accuracy (88.1%) by the human behavioural simple Bayesian model provides additional evidence for the effectiveness of the rat behavioural Bayesian model for the rat task, because the rat behavioural Bayesian model used for rats' data analysis has shown better performance than the (rat) behavioural simple Bayesian model in estimating the erroneous response patterns, at least based on the simulated data (see Section 5).

Overall, when Bayesian estimate of a hypothesis is high, the participant did use the corresponding response pattern to make their choice.

## **8.2 Human reward Bayesian model can estimate set formation for each individual**

In both rat and human 7-stage tasks, any evaluation of a participant's ED performance is based on the implicit assumption that the participant has focused on the previously relevant but currently irrelevant dimension stimuli (i.e., has formed an attentional set) before shifting their attention in the ED stage. In previous work, such set-formation is evaluated by the 'ID/ED trial difference' based on a group of participants' performance in the 7-stage task. However, forming an attentional set at *group* level does not necessarily mean that every individual participant has formed an attentional set.

The human behavioural reward Bayesian model, which considers the effect of choice feedback (in the model's prior), shows its ability to estimate whether each *individual* participant has formed an attentional set or not. If a participant has formed an attentional set before starting the ED stage, the participant would be more likely to make choices based on the response patterns linked to the irrelevant dimension at the first several trials of the ED stage. Therefore, reliably estimating whether or not a participant uses response patterns of the

irrelevant dimension at the beginning of the ED stage can be used to infer whether or not the participant has formed an attentional set. The human behavioural reward Bayesian model correctly estimated whether or not each participant used response patterns of the irrelevant dimension to make choices at the beginning of the ED stage. With this ability of the reward Bayesian model, we can accurately decide whether a participant's ED data should be used or not when analysing ED performance in any future study, or assess whether the participant is capable of forming an attentional set if they suffer brain damage of some kind or another. Note that the ability to reliably estimate whether or not a participant forms an attentional set is seen in the human behavioural reward Bayesian model, but not in the human behavioural simple Bayesian model. Therefore, although the reward Bayesian model can help analyse individual set-formation in the human 7-stage task, it is not clear whether the rat behavioural Bayesian model, which does not consider the effect of choice feedback and is an extension of the (rat) behavioural simple Bayesian model, can reliably estimate set-formation in a rat's 7-stage task.

### **8.3 Human latent probabilistic model can more reliably estimate participants' confidence than behavioural simple Bayesian model**

The human latent probabilistic model can not only help estimate erroneous response patterns used by participants and estimate whether or not participants have formed an attentional set, it might also be used to predict the participants' confidence when making choices at each trial. Specifically, the posterior estimate (i.e., b-value) of the correct (rewarded) latent hypothesis from the latent probabilistic model is highly correlated with participants' reported confidence and the correlation is higher than that from the human behavioural simple Bayesian model. The b-value from the latent probabilistic model is estimated by considering the effect of the choice feedback on learning, and there is clear evidence for the effect of feedback on participants' confidence change. Therefore, it is reasonable that the b-value from the latent probabilistic model and participants' confidence have a high correlation. In contrast, the b-value from the human behavioural simple Bayesian model is estimated without considering the effect of feedback and therefore it is reasonable to have lower correlation with participants' confidence. When participants' confidence is of interest in any set-shifting study, the high correlation between b-value and confidence could imply that the b-value could be used to infer confidence and therefore can act as a surrogate measure.

### **8.4 Human latent probabilistic model can better estimate participants' learning**

As another application of the human latent probabilistic model, the Bayesian learning

criterion based on the latent probabilistic model is consistent with participants' oral reports of the point at which they learned the correct response patterns in most stages (except for the four false positive learning cases). In comparison, the Bayesian learning criterion based on the human behavioural simple Bayesian model is not consistent with participants' oral reports in three reversal stages for most participants, and, in the ED stage, for about half of participants, and is therefore not good enough to predict when participants will have learned the correct response patterns in reversal and ED stages. Such a comparison result suggests that it is necessary to consider the effect of choice feedback in order to more accurately estimate the time point when participants have learned the correct response pattern for each learning stage. However, the Bayesian learning criterion based on the human latent probabilistic model is always satisfied when participants make six correct choices in a row at the end of each learning stage, including the four false positive learning cases. Note that the false positive learning is not due to random guesses when making choices. Instead, participants reported using esoteric response patterns, which happen to be consistent with the correct response patterns, in making their choices. These false positives could be reduced by more carefully designing and pairing blob and stick stimuli in each stage such that the rare response patterns are no longer valid to differentiate two choices at each learning trial.

On the other hand, although the Bayesian learning criterion based on the human behavioural simple Bayesian model is not consistent with participants' oral reporting, such inconsistency does not imply that the Bayesian learning criterion from the rat behavioural Bayesian model is not accurate as well. Actually, the rat behavioural Bayesian model, which considers the effect of '1st/2nd' observation at each trial, has shown better performance than the rat behavioural simple Bayesian model in estimating the rules used in a stage based on the simulated data. Of course, the rat behavioural Bayesian model (see Table 6.3 for various Bayesian models) could, and probably should, be extended to include the effect of choice feedback at each trial to further improve its ability in individual rat's data analysis, as has been shown on individual human's data analysis.

The human latent probabilistic model not only provides a better learning criterion to estimate when participants learn the correct hypothesis for each learning stage, but is also better than both the human behavioural simple Bayesian model and the frequentist approach in predicting whether or not participants are using correct response patterns to make their choice in each learning trial, including the trials before the last six correct trials in each stage. This

provides more evidence to support the view that the human latent probabilistic model, which includes the human behavioural reward Bayesian model, is a better choice to describe participants' learning at each trial.

### **8.5 Humans and rats may use different learning approaches**

While the human behavioural reward Bayesian model can effectively estimate the response patterns used by participants in the human task, it cannot estimate that the rats are using the perseveration response pattern when the rats chose bowls on the same side over six or more consecutive trials. This not only suggests that the human behavioural reward Bayesian model does not work well on the rat task, but also provides evidence that rats might use a different learning approach compared to humans in the set-shifting task. From human participants' oral reporting, I observed that humans often used a hypothesis-testing approach to learning the correct response patterns, i.e., humans first hypothesise (or guess) that a certain response pattern is the correct rule before making their choice, and then, based on the choice feedback, either continue using the same response pattern (if the choice feedback is correct) or quickly shift to another response pattern (if the choice is incorrect) in the next learning trial. This observation is consistent with previous human (Rolls et al., 1994a; Hampton et al., 2007) and rhesus monkey (Costa et al., 2015; Murray & Gaffan, 2006, in Izquierdo et al., 2017) studies which showed subjects performed reversal learning by quickly shifting rules. A recent reversal learning study with a non-Bayesian reinforcement learning approach showed that human participants implicitly *hypothesise* that a certain cue is associated with the prediction of choice outcome, and that they then either keep focusing on the cue if their prediction is correct or otherwise shift to another cue in the next trial (Akaishi et al., 2016). The human latent probabilistic model for the human task is particularly developed by considering such a hypothesis-testing learning style. However, rats may use reinforcement learning to find the correct rule in the set-shifting task, where the rats *gradually* learn the association between the correct stimulus-response rule and the reward over multiple trials. It is probably the different learning approaches between the two species that cause the poor performance of the human Bayesian analysis on the archival data from the rat task.

## 9 Human set-shifting task with Gabor patches

To further confirm the results found on the human 7-stage task with blobs and sticks as stimuli, another human 7-stage was designed by replacing the blob and stick stimuli with Gabor patches with varying orientations and frequencies (Figure 9.1). Gabor patches have the advantage that the parameters used to construct them can be considered dimensions, whereas the stick and blob stimuli used in the previous task have perceptual dimensions that can vary across individuals (e.g., as evidenced by the esoteric rules reported by some participants). Additionally, the parameters used in making Gabor patches can be adjusted to vary task difficulty and thus attentional load.

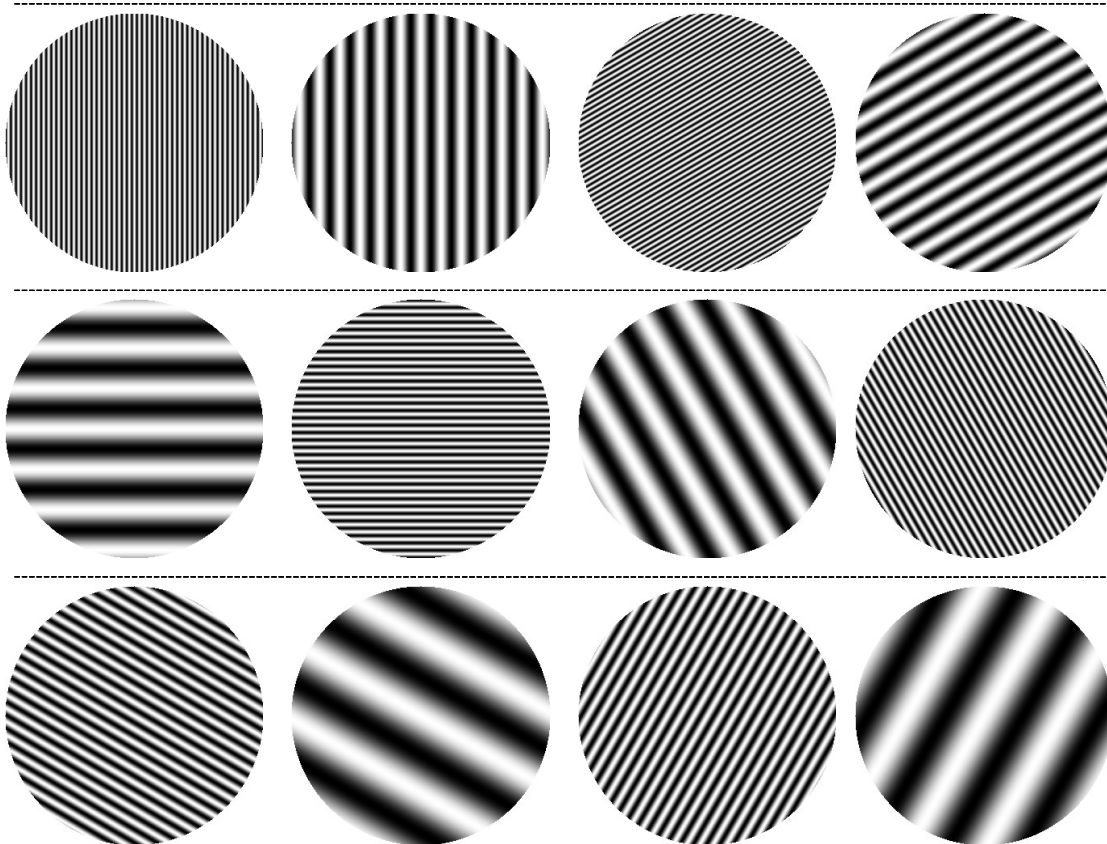


Figure 9.1: stimuli used in the human Gabor patch 7-stage task. Two dimensions of stimuli are used: orientation and frequency of lines in the Gabor patch. The four stimuli in each row were used for a unique compound stage (either CD, ID, or ED) and subsequent reversal stage.

Nineteen participants were recruited to complete the Gabor patch set-shifting task. The task data were collected and the participants' oral responses were transcribed by Meggie Rix as part of an undergraduate research project; however, I developed the software for running the task and the routines used below for analysing the resulting data. According to the criterion

used for the human blob-stick task (see Section 6.5), no participants' data were considered as outliers. The other details about the methodology are presented above in Sections 6.3-6.4.

## 9.1 Trials to criterion

A one-way between-subjects ANOVA showed that the trials to criterion (dependent variable) are significantly different across learning stages (independent variable),  $F(6, 126)=3.66$ ,  $p=0.002$ , partial  $\eta^2=0.148$ . A post-hoc Sidák-Dunn test tells us that the trials to criterion in the ED are significantly higher than those in the other five stages except for the REV1 stage, whereas there is no significant difference in the trials to criterion between any two of the six stages except for the ED stage (Figure 9.2). Note that there is no significant difference in the trials between CD and REV1, although participants often took more trials in REV1 than in CD (Figure 9.2). Similar to findings in the rat 7-stage task and in the human blob-stick 7-stage task, humans take significantly more trials to learn the ED stage compared to ID stage in the Gabor-patch 7-stage task. This indicates that, at the group level, humans have formed an attentional set before the ED stage.

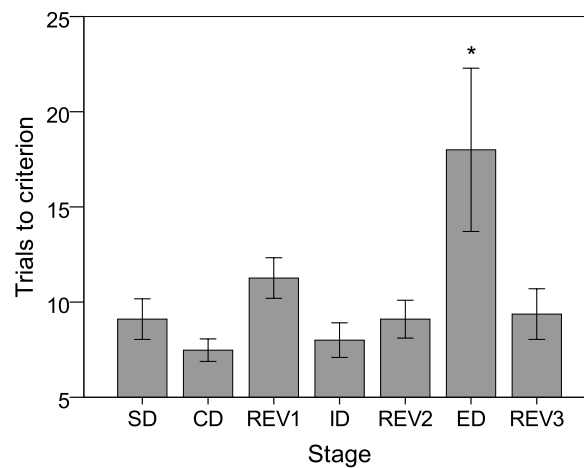


Figure 9.2: trials to criterion for each of the seven stages. Participants took significantly more trials in the ED than that in the ID stage, similar to the findings in the blob-stick human task (Figure 7.1). Vertical lines: standard errors.

## 9.2 Can the human behavioural reward Bayesian model reliably estimate which erroneous rules were used?

Just as with the human blob-stick 7-stage task, here we try to evaluate whether participants really use a specific response pattern (or *rule*) to make their choices when the Bayesian estimate of the corresponding hypothesis is high. Across all participants' seven learning stages, the human behavioural simple Bayesian model correctly estimated (see Section 7.2 for details) 76 irrelevant hypotheses and incorrectly estimated 15 irrelevant hypotheses (Table 9.1). By comparison, the human behavioural reward Bayesian model has correctly estimated

96 hypotheses and incorrectly estimated just nine hypotheses (Table 9.1, last row). The human behavioural reward Bayesian model has a trend of more correct hypothesis estimates and fewer incorrect estimates than the human behavioural simple Bayesian model, Chi-square(1) = 2.84,  $p=0.092$ .

Table 9.1: number of reward-irrelevant hypotheses correctly and incorrectly estimated by two human Bayesian models, across learning stages for all participants, with percent of correct and incorrect estimates in parentheses and expected values in italics.

	Correct hypothesis estimates	Incorrect hypothesis estimates
Human behavioural simple Bayesian model	76 (83.5%) <i>79.9</i>	15 (16.5%) <i>11.1</i>
Human behavioural reward Bayesian model	96 (91.4%) <i>92.1</i>	9 (8.6%) <i>12.9</i>

A mixed two-way ANOVA analysis showed that the numbers of estimated irrelevant/incorrect hypotheses are significantly different not only between the two Bayesian models, ( $F(1,126)=21.81$ ,  $p<0.001$ , partial  $\eta^2=0.148$ ), but also between the seven stages ( $F(6,126)=40.01$ ,  $p<0.001$ , partial  $\eta^2=0.656$ ). The interaction of model with stage is significant,  $F(6,224) = 18.12$ ,  $p<0.001$ , partial  $\eta^2=0.463$ .

Post-hoc Dunn-Sidák tests showed that significantly more irrelevant/incorrect response patterns are found in the ED and three reversal stages than in the other three stages (SD, CD, ID), and significantly fewer irrelevant/incorrect response patterns are found in the CD stage than in the SD stage, all at the 0.05 level of significance. All other comparisons were not significant (Figure 9.3).

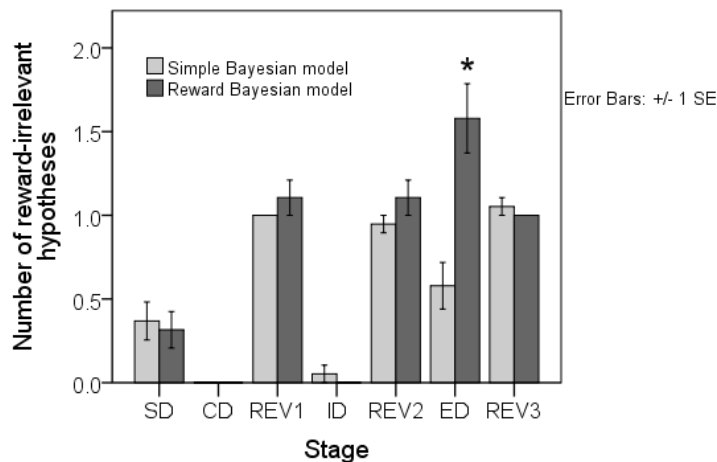


Figure 9.3: the number of irrelevant response patterns correctly estimated by the two Bayesian models for each learning stage. The human behavioural reward Bayesian model can find more irrelevant (erroneous) response patterns used by participants in the ED stage.



An additional paired-samples one-tailed t-test showed that the human behavioural reward Bayesian models found significantly more irrelevant response patterns in the ED stage than the human behavioural simple Bayesian model does (Figure 9.3),  $t(18) = 6.54$ ,  $p < 0.001$ ,  $d = 1.54$ .

### 9.3 Bayesian analysis can estimate individual set formation

The human Gabor patch 7-stage task also provides evidence to support the view that Bayesian analysis can check whether each *individual* participant has formed an attentional set. Using the human behavioural reward Bayesian model, I found that, for 17 out of 19 participants' data, the posterior probability of responding to stimuli from the previously reward relevant, but currently irrelevant, dimension becomes relatively high (i.e.,  $> 0.6$ ) at the first several trials in the ED (Figure 9.4, left). Based on the participants' oral responses, these 17 participants *did use* stimuli from the previously relevant dimension to make their choices at the beginning of the ED stage, indicating that these participants had formed an attentional set before starting the ED stage. Furthermore, for the remaining two participants' data where the posterior probability of each stimulus from the previously relevant dimension is relatively low (all less than 0.4) at the first several trials in the ED, these participants' oral responses are not consistent with making choices based on the previously relevant dimension at the beginning of the ED stage. The perfect match between the Bayesian estimate and the participant's oral response, again, suggests that the human behavioural reward Bayesian model can be used to analyse whether each individual participant has formed an attentional set or not from the first several trials in the ED stage.

In comparison, for those 17 participants who made choices based on the irrelevant dimension stimuli at the beginning of the ED stage, the human behavioural simple Bayesian model (incorrectly) estimated that 13 out of 17 participants *did not* use the stimuli in the irrelevant dimension to make their choices (Figure 9.4, right). For those two participants who did not make their choices based on irrelevant dimension stimuli at the beginning of the ED stage, the human behavioural simple Bayesian model correctly estimated that the two participants used the stimuli in the irrelevant dimension to make their choices. As explained in Section 7.3, if a participant has formed an attentional set, the participant should use stimuli from the previously relevant but currently irrelevant dimension to make their choices at the first several trials in the ED stage. This result suggests, therefore, that the human behavioural

simple Bayesian model, which did not consider the effect of choice result, cannot reliably estimate whether or not each participant has formed an attentional set.

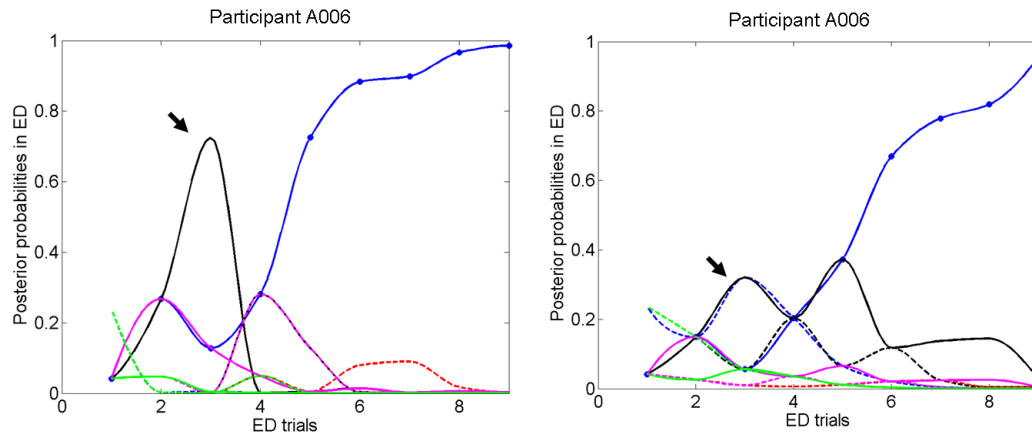


Figure 9.4: posterior probability of each hypothesis over ED trials by the two Bayesian models for one *representative* participant who made choices based on the previously relevant dimension's stimuli at the beginning (first three trials) of the ED stage, with correct estimate (arrow in the left figure) of set-formation from the human behavioural reward Bayesian model, and the incorrect estimate (arrow in the right figure) from the human behavioural simple Bayesian model. Solid and dotted black curves are for the two stimuli from the previously relevant but currently irrelevant dimension; Solid blue curve: rewarded stimulus; Dotted blue curve: non-rewarded stimulus from currently relevant dimension; Solid and dotted red curves: frequency alternation and orientation alternation hypotheses; Solid and dotted magenta and green curves: four orientation-frequency combinations.

#### 9.4 Bayesian learning criterion based on the latent probabilistic model performs better

Since we may assume that we know which response patterns (or rules) participants used for each learning trial, based on their oral report for each trial, we can evaluate how good the Bayesian learning criterion is, based on both the human behavioural simple Bayesian model and the human latent probabilistic model, and also how good the 6-in-a-row criterion is, just as in the human blob-stick 7-stage task.

From participants' oral reporting, I found that 18 out of 19 participants started to use the correct rules from one of the last six trials in each learning stage. This suggests that the 6-in-a-row criterion is well aligned with the point at which participants learned the correct rules. The remaining participant made the last six correct choices in the REV1 and REV2 stages by remembering specific exemplars.

Based on the human behavioural simple Bayesian model, I found that the Bayesian learning criterion was not satisfied by the last trial of the three reversal stages for most participants

and by the last trial of the ED stage for about half of the participants (Figure 9.5). This suggests that the human behavioural simple Bayesian model is not good enough from the perspective of learning criterion.

In comparison, with the human latent probabilistic model, the Bayesian learning criterion is satisfied by the last trial of all stages for all participants (Figure 9.6). Detailed analysis also shows that the criterion is satisfied within the last 3 trials for ID, ED, and reversal stages. This suggests that the Bayesian learning criterion based on the human latent probabilistic model is consistent with the 6-in-a-row criterion and consistent with participants' oral reporting of the point at which they learn the correct rules for all learning stages except the 2 false positive learning stages.

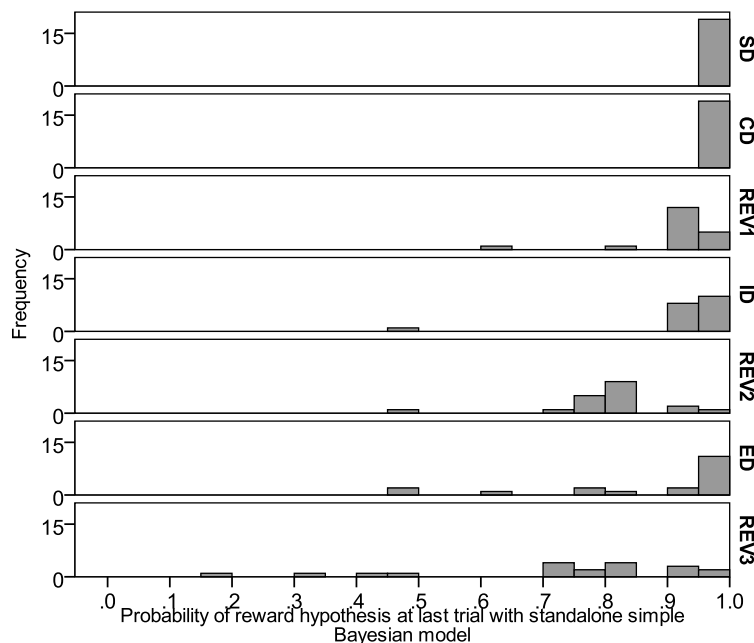


Figure 9.5: histogram of probability of 'correct'-associated hypotheses at the last trial for each learning stage over 19 participants with the human behavioural simple Bayesian model. The Bayesian learning criterion based on the simple Bayesian model is not working well, at least in the three reversal stages and the ED stage.

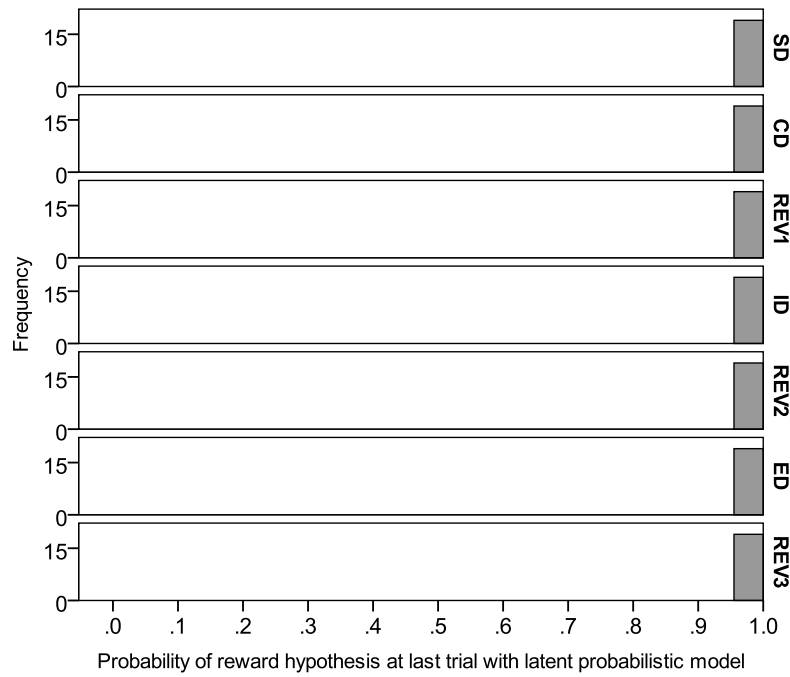


Figure 9.6: histogram of probability of ‘correct’-associated hypotheses at the last trial for each learning stage over 19 participants with the human latent probabilistic model. The Bayesian learning criterion based on the human latent probabilistic model is consistent with the 6-in-a-row criterion.

## 9.5 Discussion

The results from the Bayesian analysis of the human Gabor patch task are consistent with those from the human blob-stick task. More specifically, the analysis shows that the Bayesian estimates of both correct and erroneous response patterns match participants’ self-reports, and the set-formation of each individual participant can be reliably estimated at the beginning of the ED stage with the help of the human behavioural reward Bayesian model. Also, the analysis again shows that the Bayesian learning criterion based on the human latent probabilistic model is consistent with the 6-correct-in-a-row criterion and with the oral report from participants. All these findings again provide evidence for the usefulness of Bayesian analysis on set-shifting tasks, which indirectly provides evidence for a similar Bayesian analysis on rat data where it is impossible to know for sure the latent decision processes.

While the human Gabor patch task is an alternative to the human blob-stick task, several issues were observed in the Gabor task. First, the stimuli across stages in the Gabor task are highly related, such that the correct rule (response pattern) in one reversal stage often heavily influenced the participants’ choices in the next stage where completely new stimuli were used. For example, when participants found that the Gabor patch with more lines (stripes) was associated with the correct rule in REV1, they would naturally choose the Gabor patch with

more lines from the two new Gabor patches at the beginning of the ID stage, although the line frequencies in both new Gabor patches in the ID were *different* from those in REV1. This suggests that one rule (response pattern) in one reversal stage could naturally appear in the next stage even with completely new stimuli. However, a qualified set-shifting task design should ensure that participants would use brand new stimuli-related rules in a new stage with completely new stimuli, which is the essential objective of the ‘total change’ in the well-known CANTAB ID/ED task. The second issue from the Gabor task is that a few participants reported uncomfortable feelings (‘irritated’, ‘hate it’) when seeing the Gabor patches with either very dense (thin) lines or very few (thick) lines. In comparison, participants performing the blob-stick task did not report any such uncomfortable feelings. While participants’ self-reports may be more interpretable with the Gabor patch task because Gabor patches can be verbally described more easily than the irregular blobs and sticks in the blob-stick task, the ease of verbalising the properties of the stimuli, as shown in the self-report, might cause participants to try to solve the problem through verbal reasoning. On the other hand, Gabor patches could provide more dimensions of features, such as the stripe orientation, frequency, contrast, colour, and even movement of stripes. Therefore, the above issues from the Gabor task and obvious advantages suggest both that the human task with the Gabor patch stimuli may need to be further explored and that it could be refined for future human set-shifting study.

## **Chapter IV: Future Work and Conclusions**

### **10 Future work**

Building on the developed effective Bayesian models and interesting and novel findings on rat and human 7-stage tasks, multiple lines of future work can be investigated, including the development of a reward Bayesian model specifically for the rat task, the development of adaptive design for the 7-stage tasks, and further evaluation of the Bayesian learning criterion, particularly on the rat set-shifting task.

#### **10.1 Refinement of Bayesian model**

Currently, the reward information (i.e., ‘whether a participant got reward or not for each trial’) has been included in the latent probabilistic model for the human set-shifting task. However, the direct application of this human model to the analysis of the rat set-shifting task was not successful, even with various parameter settings in the reward model. We speculate that this might be because the reward model developed for human task is suitable for the human hypothesis-testing learning style but not suitable for rats’ associative learning style. Therefore, one future study is to investigate how to embed reward information into the Bayesian model for the rat set-shifting task. Such a refined model would not only consider which response pattern the rat is using for bowl choice, but also consider which response pattern is possibly associated with food reward, both from the experimenter’s point of view.

In addition, all the reported findings were obtained with a specific set of pre-fixed parameter values in the rat model and the human model respectively. It is worth investigating whether we would still obtain similar findings by varying the parameter values in both the rat model and human model. Initial Bayesian re-analysis for some of the key findings has already shown that both rat and human models are robust with varying model parameter values (Appendix A). More comprehensive investigation of model parameters could be performed on all the reported findings in the future.

#### **10.2 Adaptive design**

Traditionally, for each trial of a learning stage in the 7-stage task, researchers pre-determine which stimuli will be put into each bowl. Such a fixed trial design is meant to counterbalance across rats the positioning of different stimuli, such that the reward-associated stimulus (e.g.,

M1) appears both in the left and in the right bowls over trials, and the pairing of different stimuli (e.g., O1 and O2) of the other dimension to the correct stimulus. However, this fixed design does not consider the performance of each rat (or human) over trials. For example, a rat may be using an incorrect response pattern to choose bowls and the food reward may happen to appear in the chosen bowls in several consecutive trials. In this case, the rat may strengthen the association between the incorrect response pattern and food reward. Consequently, the rat has to spend more trials in finding the food-associated response pattern.

Using Bayesian analysis of a rat's performance in the previous trial, we can adaptively determine which bowl will contain which stimuli for a given trial. I call such a dynamic process *adaptive trial design*. Different goals of adaptive trial design require different design principles. One possible goal of adaptive trial design is to potentially speed up rats' learning in each stage. This goal could be achieved by minimising ambiguity and maximising information available to the rat concerning its most likely hypothesis. For example, suppose Bayesian analysis tells us that the rat is likely to respond to O2, even though M1 is reward-associated, then for the next trial, M1 and O2 can be placed in different bowls. In this case, the rat would not be rewarded if it chooses the bowl with O2. This will have the effect of decreasing the association between the incorrect response pattern ('O2') and food reward. In this way, the rat may more quickly find the correct response pattern.

The above example also indicates the principle of adaptive design for the goals of speeding up rats' learning: put the food-associated stimulus into the bowl that the rat is least likely to choose. In more detail, suppose M1 is associated with food reward, each trial can be adaptively designed as follows:

- (1) Determination of stimulus pairs: if the Bayesian probability of O2 is higher than that of O1 based on the last trial's observation, then put M1 and O1 into one bowl, and put the remaining stimuli (M2 and O2) into the other bowl. Otherwise, put M1 and O2 into one bowl.
- (2) Determination of side: If the rat is more likely to choose the left bowl than the right bowl based on Bayesian analysis of all four spatial hypotheses in the last trial, then put the food-associated M1 (and its paired odour stimulus) into the right bowl. Otherwise, put M1 into the left bowl.

Based on the above principles, we can adaptively design each trial based on the Bayesian analysis of previous trials, which will potentially speed up rats' learning process.

Figure 10.1 shows the brief structure of the adaptive design software implemented in MATLAB. For each trial, the software displays the bowl information through the user interface (see the 'Left' and 'Right' bowls in Figure 10.2). Based on the displayed bowl design, the user (i.e., experimenter) sets up the real trial, and then observes and inputs the rat's behaviour through the user interface (Figure 10.2). Then the software performs Bayesian analysis of the current trial to estimate the posterior probability of each hypothesis. Based on the posterior probabilities and the adaptive design principles, the software determines and displays which bowl should contain which stimuli for the next trial. The above process is repeated until each learning stage is finished by satisfying the Bayesian criterion (i.e., Bayesian probability of the correct hypothesis is larger than 0.95 by default). The rat's behavioural responses and the estimated posterior probabilities for each hypothesis are saved into one Excel file for each rat.

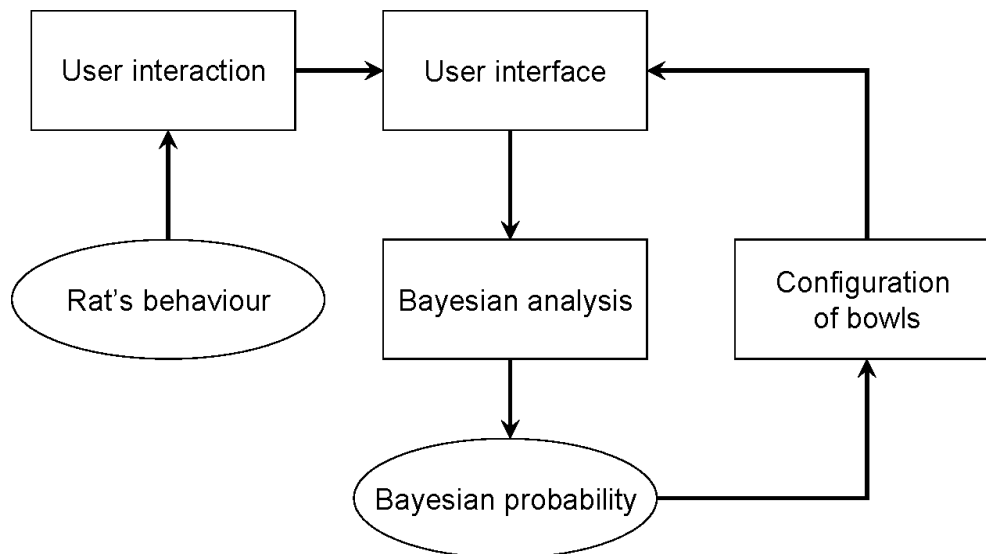


Figure 10.1: the basic framework of adaptive design.



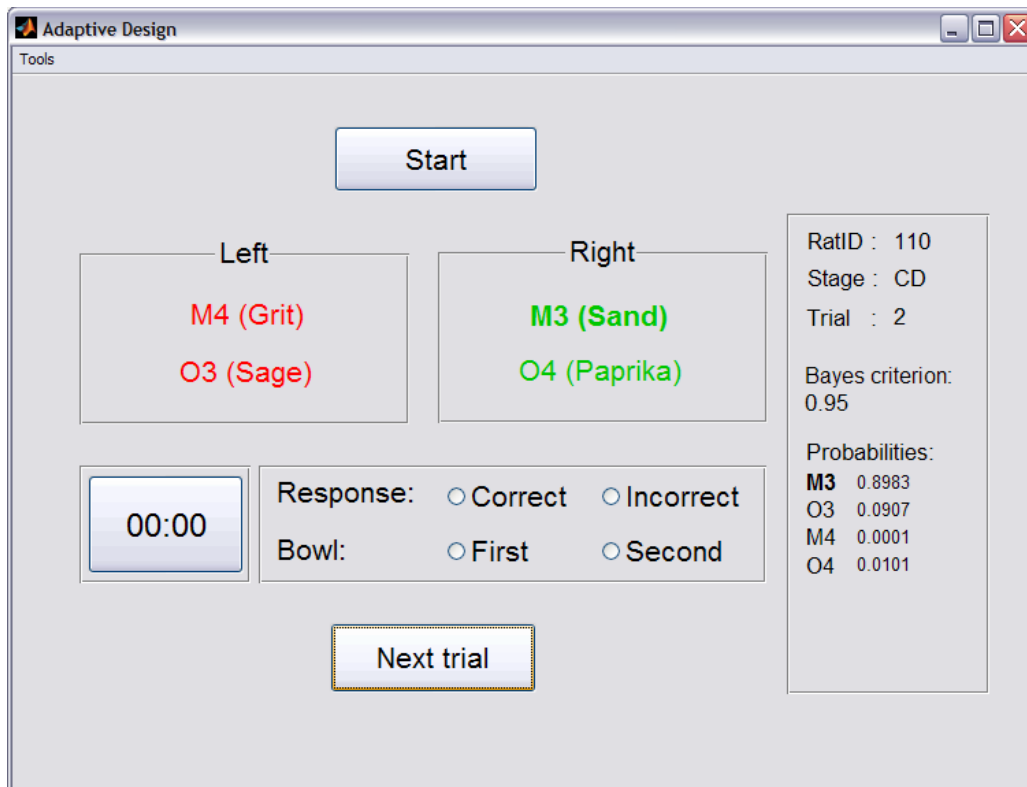


Figure 10.2: a screen-shot of the adaptive design software. The stimuli in the ‘Left’ and the ‘Right’ bowls are determined from the adaptive design process, with the bold green stimulus associated with reward. User can input rat’s behaviour by clicking one of the two bowls, according to whether rat’s response is correct (i.e., got reward) or not and whether it digs the 1<sup>st</sup> or 2<sup>nd</sup> bowl it encounters, and recording the time spent to choose a bowl. Clicking ‘Next trial’ will trigger the Bayesian analysis of the current trial and then adaptively determine and display the pairing of stimuli in each bowl for the next trial. The right column on the user interface displays general information, including rat identity, current stage and trial number, the Bayesian learning criterion, and the estimated posterior probabilities of the four perceptual hypotheses from the previous trial.

Using the adaptive design software, 12 normal rats and 11 mPFC-lesioned rats have performed the 7-stage task with each trial adaptively designed. The initial analysis of these 23 rats’ data showed that both normal and lesioned rats did not show a difference in the number of trials between ID and ED stages. It seems that the adaptive design made both control and lesioned rats learn faster in the ED stage by helping rats disambiguate the relationship between stimuli and reward, such that the ID/ED difference observed in previous studies disappeared. However, considering the following factors, it is too early to draw any convincing conclusion from the initial data analysis:

- Bayesian model issue: the rat behavioural Bayesian model has not been well developed to capture all relevant observation information, particularly the reward observation at

each trial. The current incomplete Bayesian model may not be able to well predict what hypothesis a rat would use to choose bowls in next trials, and therefore make the adaptive design of learning trials inappropriate. Refinement of the current Bayesian model could solve the ID/ED non-difference issue.

- Data collection issue: the data collection procedure may not be well designed. Specifically, for each of the first four trials in each learning stage of the 7-stage task, the rat was allowed to dig in the other bowl if it dug first in the unrewarded bowl. This special step was designed to make the rat establish what the stimuli were in the bowls, and implicitly sped up rats' learning. However, it may also suppress the perseveration errors that the rats could potentially have made. From this perspective, the lack of difference between the ID and ED trials based on adaptive design may be potentially from the special data collection step during the first four trials in the ED stage. To avoid the potential effect of this special step, new rats' data should be collected and analysed using the adaptive design software with the exclusion of this special step.

Therefore, more study is necessary to explore the adaptive design for attentional set-shifting tasks.

In addition, besides the above goal and principle of adaptive design, we may also explore other design goals. An alternative goal of adaptive design is to *maximally distinguish the likelihoods of the competing hypotheses* (Liepe et al., 2013). To achieve this goal, the adaptive design should take into account the posteriors from the previous trial and search for the configuration of bowls that, on average across all possible responses to the configuration, results in the highest variability in the posterior probabilities after the current trial. In essence, what the adaptive design typically aims for is a posterior probability of 1 for the correct hypothesis and 0 for all remaining hypotheses. The only concern is whether such an adaptive trial design would work if the conditions are changing, as we expect during learning in each stage of the 7-stage task. One future goal is to explore which design goal and corresponding design principle is more appropriate for the attentional set-shifting task.

### **10.3 Bayesian learning criterion with fixed design**

Besides the refinement of the Bayesian model for the rat task, and adaptive design, another future goal is to further evaluate the application of the Bayesian learning criterion in the traditional fixed design of bowls. One clear prediction from the current work is that if the

rats' behaviour is judged by a Bayesian criterion, then the seemingly large variability among the learning rates of rats might be more homogeneous than when the n-in-a-row criterion is used. If so, it could enhance statistical power and thereby reduce the numbers of rats required for research, both of which are desirable aims. Thus, simply comparing the apparent variability of two cohorts of rats, one of which is trained using the previous 6-in-row criterion and the other trained using the Bayesian learning criterion, would seem to be a reasonable test of the relative utility of the Bayesian approach.

## 11 Conclusions

Attentional set-shifting tasks in both animals and humans have been used extensively to study cognitive flexibility. Multiple behavioural learning criteria (e.g., 6-correct-choices-in-a-row, 10-out-of-12-correct-choices) have been adopted in attentional set-shifting tasks. All the learning criteria take a classical null-vs-alternative hypothesis inferential testing approach, where the null hypothesis is that subjects use random guesses to make their choices during learning and the alternative hypothesis is that subjects use the correct rule to make their choices. The problem is that, besides the random guess and the correct rule, there are multiple alternative erroneous rules that subjects could use to make choices. Before establishing the correct stimulus-reward association, a rat (similarly for human participants) may have tried other non-random, but reward-irrelevant, spatial patterns or stimulus characteristics to make choices. The traditional hypothesis-testing approach can only help us to determine the point at which subjects have learned, but it does not allow us to determine which erroneous response patterns or rules (either perceptual or spatial) could have been tried in each learning stage by each subject. As a result, the data collected from set-shifting tasks typically can only be analysed at the group level, e.g., comparing whether two groups of subjects have similar performance in the ED stage or a reversal stage.

To solve the issues and limitations of the traditional frequentist hypothesis-testing approach, I developed a novel Bayesian approach initially for the rat set-shifting task and then further extended the model for human tasks. Bayesian analysis of individual rats' learning behaviour provides us with detailed information about what response patterns (or rules) may have been tried in each learning trial for each rat. By comparing the Bayesian probability of the reward-associated hypothesis with the pre-set high threshold value (0.95) in each learning trial, we can determine whether or not each rat has learned the stimulus-reward association in a timely manner for each learning stage. Such a Bayesian learning criterion is theoretically better than frequentist learning criteria (e.g., 6-correct-choice-in-a-row), because the Bayesian approach has estimated the probability of all pre-established spatial and perceptual hypotheses when deciding to accept the reward-relevant hypothesis.

Bayesian analysis provides a potentially more powerful approach than the frequentist approach to analysing the data collected by set-shifting tasks. Based on the individual analysis with the developed rat behavioural Bayesian model, I found that all rats tried various

spatial patterns of responding while they were learning the stimulus-reward associations, and mPFC-lesioned rats had more spatial trials than control rats in the ED and reversal stages. It is difficult, if not impossible, to find such results using only the frequentist approach.

All the Bayesian analysis results from rat set-shifting tasks are valid only if the Bayesian estimates of the hypotheses can really suggest the possible response patterns that rats are actually using for bowl choice. Considering that we never know for certain what exact response pattern a given rat uses to choose bowls at each trial, the validity of the rat behavioural Bayesian model was firstly investigated by Bayesian analysis on simulated data where the ground-truth response patterns underlying the simulated data were known. The analysis on the simulated data showed that, when the Bayesian estimate of a specific hypothesis is high in a learning stage the (virtual) rat did use the hypothesis to make its choice in the stage.

To further support the validity of the Bayesian analysis, I implemented an analogous human 7-stage task and purposely collected human participants' oral reports on which hypotheses they actually used to make their choice for each learning trial. In view of the possible different characteristics between humans and rats in learning, the rat behavioural Bayesian model developed for the rat task cannot be directly applied to the human task. Instead of just using the human behavioural simple Bayesian model, which is a simplified version of the rat behavioural Bayesian model, I also developed a new, human latent probabilistic model (including a human behavioural reward Bayesian model) which considers the effect of choice feedback on learning. Bayesian analysis of the human task with both the human behavioural simple Bayesian model and the reward Bayesian model clearly shows that the Bayesian estimate of both reward-relevant and reward-irrelevant (i.e., erroneous) hypotheses match what participants orally reported. This provides supportive, albeit indirect, evidence for the validity of performing Bayesian analysis on rats' data, as the rat behavioural Bayesian model for the rat task is an advanced version of the human behavioural simple Bayesian model for the human task.

In addition, the strong correspondence between participants' oral reporting and the Bayesian estimate of perceptual hypotheses from the previously rewarded but currently irrelevant dimension at the beginning of the ED stage also suggests that the human behavioural reward Bayesian model can also help decide whether or not each participant has formed an

attentional set before shifting to the newly relevant dimension. This provides a novel way to either accept or discard a participant's data when collecting data for any subsequent analysis of ED performance. In contrast, the traditional 'ID/ED difference' method is used to decide whether a group of participants, but not each individual participant, have formed an attentional set or not. Therefore, the reward Bayesian model can help improve the power of data analysis by excluding those data where an attentional set has not been well formed, or can help classify data such that further analysis can be performed in the group of participants who fail to form an attentional set.

In conclusion, I developed two probabilistic (Bayesian) models which can effectively and reliably analyse subjects' discrimination learning at the individual level. The models not only provide an alternative learning criterion to decide when subjects have learned the correct rules, but also have helped find what erroneous rules may have been tried by subjects before learning the correct rules. The model for the human task can also help decide whether subjects have a well-formed attentional set before set-shifting, thereby improving the power of data analysis. Of course, this is only the first step to the successful application of a Bayesian approach to set-shifting tasks. The Bayesian models could be further refined, applied to adaptive design and other set-shifting tasks (e.g., 4ID), and eventually applied to the study of more general cognitive flexibility and learning.

## References

- Abdul-Monim, Z., Reynolds, G. P., & Neill, J. C. (2006). The effect of atypical and classical antipsychotics on sub-chronic PCP-induced cognitive deficits in a reversal-learning paradigm. *Behav. Brain Res.*, 169, 263-273.
- Akaishi, R., Kolling, N., Brown, J. W., & Rushworth, M. (2016). Neural mechanisms of credit assignment in a multicue environment. *The Journal of Neuroscience*, 36(4), 1096-1112.
- Allison, C., & Shoaib, M. (2013). Nicotine improves performance in an attentional set shifting task in rats. *Neuropharmacology*, 64, 314-320.
- Aloi, M., Rania, M., Caroleo, M., Bruni, A., Palmieri, A., Cauteruccio, M., De Fazio, P., & Segura-Garcia, C. (2015). Decision making, central coherence and set-shifting: a comparison between binge eating disorder, anorexia nervosa and healthy controls. *BMC Psychiatry*, 15(6).
- Aoki, S., Andrew, X., Liu, W., Zucca, A., Zucca, S., & Wickens, J. R. (2015). Role of striatal cholinergic interneurons in set-shifting in the rat. *The Journal of Neuroscience*, 35(25), 9424-9431.
- Beas, B. S., McQuail, J. A., Banuelos, C., Setlow, B., Bizon, J. L. (2017). Prefrontal cortical GABAergic signaling and impaired behavioral flexibility in aged F344 rats. *Neuroscience*, 345, 274-286.
- Beas, B. S., Setlow, B., & Bizon, J. L. (2013). Distinct manifestations of executive dysfunction in aged rats. *Neurobiol. Aging*, 34(9), 2164-2174.
- Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *J. Gen. Psychol.*, 39, 15-22.
- Birrell, J. M., & Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *J. Neurosci.*, 20, 4320-4324.
- Bissonette, G. B., Bae, M. H., Suresh, T., Jaffe, D. E., & Powell, E.M. (2010). Astrocyte-mediated hepatocyte growth factor/scatter factor supplementation restores GABAergic interneurons and corrects reversal learning deficits in mice. *J. Neurosci.*, 30, 2918-2923.
- Bissonette, G. B., Martins, G. J., Franz, T. M., Harper, E. S., Schoenbaum, G., & Powell, E. M. (2008). Double dissociation of the effects of medial and orbital prefrontal cortical lesions on attentional and affective shifts in mice. *J. Neurosci.*, 28, 11124-11130.
- Bissonette, G. B., & Powell, E. M. (2012). Reversal learning and attentional set-shifting in mice. *Neuropharmacology*, 62, 1168-1174.
- Bissonette, G. B., Powell, E. M., & Roesch, M. R. (2013). Neural structures underlying set-shifting: Roles of medial prefrontal cortex and anterior cingulate cortex. *Behav. Brain Res.*, 250, 91-101.
- Bissonette, G. B., Schoenbaum, G., Roesch, M. R., & Powell, E. M. (2015). Interneurons are necessary for coordinated activity during reversal learning in orbitofrontal cortex. *Biol Psychiatry*, 77(5), 454-464.
- Block, A. E., Dhanji, H., Thompson-Tardif, S. F., & Floresco, S. B. (2007). Thalamic-prefrontal cortical-ventral striatal circuitry mediates dissociable components of strategy set shifting. *Cereb. Cortex*, 17, 1625-1636.
- Bondi, C. O., Rodriguez, G., Gould, G. G., Frazer, A., & Morilak, D. A. (2008). Chronic unpredictable stress induces a cognitive deficit and anxiety-like behavior in rats that is prevented by chronic antidepressant drug treatment. *Neuropsychopharmacology*, 33, 320-331.
- Borkowska, A., Rybakowski, J. K. (2001). Neuropsychological frontal lobe tests

- indicate that bipolar depressed patients are more impaired than unipolar. *Bipolar Disord*, 3(2), 88-94.
- Boulougouris, V., Dalley, J. W., & Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behav. Brain Res.*, 179, 219-28.
- Brady, A. M., & Floresco, S. B. (2015). Operant procedures for assessing behavioral flexibility in rats. *Journal of Visualized Experiments*, 96, e52387.
- Brigman, J. L., Bussey, T. J., Saksida, L. M., & Rothblat, L. A. (2005). Discrimination of multidimensional visual stimuli by mice: intra- and extradimensional shifts. *Behav. Neurosci.*, 119, 839-842.
- Brigman, J. L., Daut, R. A., Wright, T., Gunduz-Cinar, O., Graybeal, C., & Davis, M. I., et al. (2013). GluN2B in corticostriatal circuits governs choice learning and choice shifting. *Nat. Neurosci.*, 16, 1101-1110.
- Brigman, J. L., Mathur, P., Harvey-White, J., Izquierdo, A., Saksida, L. M., Bussey, T. J., Fox, S., Deneris, E., Murphy, D. L., & Holmes, A. (2010). Pharmacological or genetic inactivation of the serotonin transporter improves reversal learning in mice. *Cereb Cortex*. 20, 1955-1963.
- Brockett, A. T., LaMarca, E. A., & Gould, E. (2015). Physical exercise enhances cognitive flexibility as well as astrocytic and synaptic markers in the medial prefrontal cortex. *PLoS ONE*, 10(5), e0124859.
- Brooks, S.P., Janghra, N., Higgs, G. V., Bayram-Weston, Z., Heuer, A., Jones, L., & Dunnett, S. B. (2012). Selective cognitive impairment in the YAC128 Huntington's disease mouse, *Brain Res. Bull.*, 88, 121-129.
- Brown, V. J., & Bowman, E. M. (2002). Rodent models of prefrontal cortical function. *Trends in Neurosciences*. 25(7), 340-343.
- Brown, V. J., & Tait, D. S. (2016). Attentional Set-Shifting Across Species, *Curr Topics Behav Neurosci.*, DOI 10.1007/7854\_2015\_5002
- Brushfield, A. M., Luu, T. T., Callahan, B. D., & Gilbert, P. E. (2008). A comparison of discrimination and reversal learning for olfactory and visual stimuli in aged rats. *Behavioral Neuroscience*. 122, 54-62.
- Chamberlain, S. R., Muller, U., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2006). Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science*, 311, 861-863.
- Chase, E. A., Tait, D. S., & Brown, V. J. (2012). Lesions of the orbital prefrontal cortex impair the formation of attentional set in rats. *Eur. J. Neurosci.*, 36, 2368-2375.
- Chau, B. K., Sallet, J., Papageorgiou, G. K., Noonan, M. P., Bell, A. H., & Walton, M. E., et al. (2015) Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. *Neuron*, 87, 1106-1118.
- Cho, K. K. A., Hoch, R., Lee, A. T., Patel, T., Rubenstein, J. L. R. , & Sohal, V. S. (2015). Gamma rhythms link prefrontal interneuron dysfunction with cognitive inflexibility in *Dlx5/6*<sup>+/-</sup> mice. *Neuron*, 85(6), 1332-1343.
- Chudasama, Y., & Robbins, T. W. (2006). Functions of frontostriatal systems in cognition: Comparative neuropsychopharmacological studies in rats, monkeys and humans. *Biological Psychology*, 73 (1), 19-38.
- Clarke, H. F., Hill, G. J., Robbins, T. W., & Roberts, A. C. (2011). Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. *J. Neurosci.*, 31, 4290-4297.
- Clarke, H. F., Walker, S. C., Crofts, H. S., Dalley, J. W., Robbins, T. W., & Roberts, A. C. (2005). Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *J. Neurosci.*, 25, 532-538.



- Clarke, H. F., Walker, S. C., Dalley, J. W., Robbins, T. W., & Roberts, A. (2007). Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cerebral Cortex*, 17(1), 18-27.
- Cools, R. (2006). Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neurosci. Biobehav. Rev.*, 30, 1-23.
- Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2015). Reversal learning and dopamine: a Bayesian perspective. *The Journal of Neuroscience*, 35(6), 2407-2416.
- Crofts, H. S., Dalley, J. W., Collins, P., Van Denderen, J. C. M., Everitt, B. J., et al. (2001). Differential effects of 6-OHDA lesions of the frontal cortex and caudate nucleus on the ability to acquire attentional set. *Cereb. Cortex*, 11, 1015-1026.
- Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: implications for clinical and developmental neuroscience. *Trends Neurosci.*, 38, 571-578.
- Dalton, G. L., Ma, L. M., Phillips, A. G., & Floresco, S. B. (2011). Blockade of NMDA GluN2B receptors selectively impairs behavioral flexibility but not initial discrimination learning. *Psychopharmacology*, 216, 525-535.
- Das, A., Dikshit, M., & Nath, C. (2001). Profile of acetylcholinesterase in brain areas of male and female rats of adult and old age. *Life Sci.*, 68, 1545-1555.
- Desai, S. J., Allman, B. L., Rajakumar, N. (2017). Combination of behaviorally sub-effective doses of glutamate NMDA and dopamine D1 receptor antagonists impairs executive function. *Behav Brain Res.*, 323, 24-31.
- De Steno, D. A., & Schmauss, C. (2009). A role for dopamine D2 receptors in reversal learning. *Neuroscience*, 162, 118-127.
- Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Primate analogue of the Wisconsin Card Sorting Test: effects of excitotoxic lesions of the prefrontal cortex in the marmoset. *Behav. Neurosci.*, 110, 872-886.
- Dias, R., Robbins, T. W., & Roberts, A. C. (1997) Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. *J Neurosci*, 17(23), 9285-9297.
- Dickson, P. E., Cairns, J., Goldowitz, D., & Mittleman, G. (2017). Cerebellar contribution to higher and lower order rule learning and cognitive flexibility in mice. *Neuroscience*, 345, 99-109.
- Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., Goldman, D., & Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc. Natl. Acad. Sci. U S A*, 98, 6917-6922.
- Elliott, R., McKenna, P. J., Robbins, T. W., & Sahakian, B. J. (1995). Neuropsychological evidence for frontostriatal dysfunction in schizophrenia. *Psychol Med.*, 25, 619-630.
- Fazekas, I., Karacsony, Z., & Libor, Z. (2010). Longest runs in coin tossing. Comparison of recursive formulae, asymptotic theorems, computer simulations. *Acta Univ. Sapientiae, Mathematica* 2, 2, 215-228.
- Fineberg, N. A., Day, G. A., de Koenigswarter, N., Reghunandanan, S., Kolli, S., Jefferies-Sewell, K., Hranov, G., & Laws, K. R. (2015). The neuropsychology of obsessive-compulsive personality disorder: a new analysis. *CNS Spectr.*, 20, 490-499.
- Fisher, R.A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron*, 1, 3-32.
- Floresco, S. B. (2013). Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. *Front Neurosci.*, 7:62.
- Floresco, S. B., Block, A. E., & Tse, M. T. (2008). Inactivation of the medial

- prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behav Brain Res.*, 190, 85-96.
- Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., & Tse, M. T. (2006). Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology*, 31, 297-309.
- Floresco, S. B., & Jentsch, J. D. (2011). Pharmacological enhancement of memory and executive functioning in laboratory animals. *Neuropsychopharmacology* 36, 227-250.
- Floresco, S. B., Zhang, Y., & Enomoto, T. (2009). Neural circuits subserving behavioural flexibility and their relevance to schizophrenia. *Behav Brain Res*, 204, 396–409.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2003). *Bayesian data analysis*. Second Edition, CRC Press.
- Geva-Sagiv, M., Las, L., Yovel, Y., & Ulanovsky, N. (2015). Spatial cognition in bats and rats: from sensory acquisition to multiscale maps and navigation. *Nat Rev Neurosci*, 16, 94-108.
- Ghods-Sharifi, S., Haluk, D. M., & Floresco, S. B. (2008). Differential effects of inactivation of the orbitofrontal cortex on strategy set-shifting and reversal learning. *Neurobiol. Learn. Mem.*, 89, 567-73.
- Glickstein, S. B., Desteno, D. A., Hof, P. R., & Schmauss, C. (2005). Mice lacking dopamine D2 and D3 receptors exhibit differential activation of prefrontal cortical neurons during tasks requiring attention. *Cereb. Cortex*, 15, 1016-1024.
- Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. USA*, 108(3), 15647-15654.
- Graybeal, C., Bachu, M., Mozhui, K., Saksida, L. M., Bussey, T. J., & Sagalyn, E., et al. (2014). Strains and stressors: an analysis of touchscreen learning in genetically diverse mouse strains. *PLoS ONE*, 9, e87745.
- Groman, S. M., James, A. S., Seu, E., Crawford, M. A., Harpster, S. N., & Jentsch, J. D. (2013). Monoamine levels within the orbitofrontal cortex and putamen interact to predict reversal learning performance. *Biol. Psychiatry*, 73, 756-762.
- Haluk, D. M., & Floresco, S. B. (2009). Ventral striatal dopamine modulation of different forms of behavioral flexibility. *Neuropsychopharmacology*, 34, 2041-2052.
- Hamilton, D. A., & Brigman, J. L. (2015). Behavioral flexibility in rats and mice: contributions of distinct frontocortical regions. *Genes, Brain and Behavior*, 14, 4-21.
- Hampton, A. N, Adolphs, R., Tyszka, M. J., & O'Doherty, J. P. (2007). Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron*, 55, 545-555.
- Harrison, P.J. (1999). The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain*, 122, 593-624.
- Hernandez, A. R., Reasor, J. E., Truckenbrod, L. M., Lubke, K. N., Johnson, S. A., Bizon, J. L., Maurer, A. P., Burke, S. N. (2017). Medial prefrontal-perirhinal cortical communication is necessary for flexible response selection. *Neurobiology of Learning and Memory*, 137, 36-47.
- Heyser, C. J., Fienberg, A. A., Greengard, P., & Gold, L. H. (2000). DARPP-32 knockout mice exhibit impaired reversal learning in a discriminated operant task. *Brain Res.*, 867, 122-130.
- Idris, N. F., Neill, J. C., Grayson, B., Bang-Andersen, B., Witten, L. M., Brennum, L. T., et al. (2010). Sertindole improves subchronic PCP-induced reversal learning and episodic memory deficits in rodents: involvement of 5-HT6 and 5-HT2A receptor mechanisms. *Psychopharmacology*, 208, 23-36.

- Izquierdo, A., Belcher, A. M., Scott, L., Cazares, V. A., Chen, J., O'Dell, S. J., Malvaez, M., Wu, T., & Marshall, J. F. (2010). Reversal-specific learning impairments after a binge regimen of methamphetamine in rats: possible involvement of striatal dopamine. *Neuropsychopharmacology*, 35, 505-514.
- Izquierdo, A., Darling, C., Manos, N., Pozos, H., Kimm C., & Ostrander, S., et al. (2013). Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. *J. Neurosci.*, 33, 4105-4109.
- Izquierdo, A., & Jentsch, J. D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology*, 219, 607-620.
- Janitzky, K., Lippert, M. T., Engelhorn, A., Tegtmeier, J., Goldschmidt, J., Heinze, H.-J., & Ohl, F. W. (2015). Optogenetic silencing of locus coeruleus activity in mice impairs cognitive flexibility in an attentional set-shifting task. *Frontiers in Behavioral Neuroscience*, 9:286.
- Jazbec, S., Pantelis, C., Robbins, T., Weickert, T., Weinberger, D. R., & Gldberg, T. E. (2007). Intra-dimensional/extra-dimensional set-shifting performance in schizophrenia: impact of distractors. *Schizophr Res.*, 89, 339-349.
- Jeevakumar, V., Driskill, C., Paine, A., Sobhanian, M., Vakil, H., Morris, B., Ramos, J., Kroener, S. (2015). Ketamine administration during the second postnatal week induces enduring schizophrenia-like behavioral symptoms and reduces parvalbumin expression in the medial prefrontal cortex of adult mice. *Behav. Brain Res.*, 282, 165-175.
- Jentsch, J.D., & Roth, R.H. (1999). The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology*, 20, 201-225.
- Jett, J. D., Boley, A. M., Girotti, M., Shah, A., Lodge, D. J., & Morilak, D. A. (2015). Antidepressant-like cognitive and behavioral effects of acute ketamine administration associated with plasticity in the ventral hippocampus to medial prefrontal cortex pathway. *Psychopharmacology*, 232, 3123-3133.
- Jett, J. D., Bulin, S. E., Hatherall, L. C., McCartney, C. M., & Morilak, D. A. (2017). Deficits in cognitive flexibility induced by chronic unpredictable stress are associated with impaired glutamate neurotransmission in the rat medial prefrontal cortex. *Neuroscience*, 346, 284-297.
- Kehagia, A. A., Barker, R. A., & Robbins T. W. (2010). Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. *Lancet Neurol.* 9, 1200-1213.
- Kim, D. H., Choi, B. R., Jeon, W. K., & Han, J. S. (2016). Impairment of intradimensional shift in an attentional set-shifting task in rats with chronic bilateral common carotid artery occlusion. *Behavioural Brain Research*, 296, 169-176.
- Klanker, M., Feenstra, M., & Denys, D. (2013). Dopaminergic control of cognitive flexibility in humans and animals. *Front Neurosci.*, 7:201.
- Klanker, M., Sandberg, T., Joosten, R., Willuhn, I., Feenstra, M., & Denys, D. (2015). Phasic dopamine release induced by positive feedback predicts individual differences in reversal learning. *Neurobiol. Learn. Mem.*, 125, 135-145.
- Kruzich, P. J., & Grandy, D. K. (2004). Dopamine D2 receptors mediate two-odor discrimination and reversal learning in C57BL/6 mice. *BMC Neurosci.*, 5:12.
- Kruzich, P. J., Mitchell, S. H., Younkin, A., & Grandy, D. K. (2006). Dopamine D2 receptors mediate reversal learning in male C57BL/6J mice. *Cogn. Affect. Behav. Neurosci.*, 6, 86-90.
- Lange, F., Seer, C., Loens, S., Wegner, F., Schrader, C., Dressler, D., Dengler, R., & Kopp, B. (2016). Neural mechanisms underlying cognitive inflexibility in Parkinson's disease. *Neuropsychologia*, 93, 142-150.

- Lapiz, M. D., Bondi, C. O., & Morilak, D. A. (2007). Chronic treatment with desipramine improves cognitive performance of rats in an attentional set-shifting test. *Neuropsychopharmacology*, 32, 1000-1010.
- Lapiz, M. D., & Morilak, D. A. (2006). Noradrenergic modulation of cognitive function in rat medial prefrontal cortex as measured by attentional set shifting capability. *Neuroscience*, 137, 1039-1049.
- Lawrence, A. D., Sahakian, B. J., Hodges, J. R., Rosser, A. E., Lange, K. W., & Robbins, T. W. (1996). Executive and mnemonic functions in early Huntington's disease. *Brain*, 119, 1633-1645.
- Lawrence, A. D., Sahakian, B. J., Rogers, R. D., Hodge, J. R., & Robbins, T. W. (1999). Discrimination, reversal, and shift learning in Huntington's disease: mechanisms of impaired response selection. *Neuropsychologia*, 37(12), 1359-1374.
- Lee, I., & Byeon, J. S. (2014). Learning-dependent changes in the neuronal correlates of response inhibition in the prefrontal cortex and hippocampus. *Experimental Neurobiology*, 23, 178-189.
- Leeson, V. C., Robbins, T. W., Matheson, E., Hutton, S. B., Ron, M. A., Barnes, T. R., & Joyce, E. M. (2009). Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: stability over six years and specific associations with medication type and disorganization syndrome. *Biol. Psychiatry*, 66, 586-593.
- Liepe, J., Filippi, S., Komorowski, M., & Stumpf, M. P. H. (2013). Maximizing the Information Content of Experiments in Systems Biology. *PLOS Computational Biology*, 9(1), e1002888, doi:10.1371/journal.pcbi.1002888
- Lindgren, H. S., Wickens, R., Tait, D. S., Brown, V. J., & Dunnett, S. B. (2013). Lesions of the dorsomedial striatum impair formation of attentional set in rats. *Neuropharmacology*, 71, 148-153.
- Linley, S. B., Gallo, M. M., & Vertes, R. P. (2016). Lesions of the ventral midline thalamus produce deficits in reversal learning and attention on an odor texture set shifting task. *Brain Research*, 1649, 110-122.
- Lloyd, K., & Leslie, D. S. (2013). Context-dependent decision-making: a simple Bayesian model. *Journal of the Royal Society Interface*, 10(82), doi:10.1098/rsif.2013.0069
- Mackintosh, N. J. (1974). *The psychology of animal learning*. New York, NY: Academic Press.
- Malá, H., Andersen, L. G., Christensen, R. F., Felbinger, A., Hagstrøm, J., Meder, D., & Mogensen, J. (2015). Prefrontal cortex and hippocampus in behavioural flexibility and posttraumatic functional recovery: reversal learning and set-shifting in rats. *Brain Research Bulletin*, 116, 34-44.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, 125(3), 624-639.
- Mar, A. C., Horner, A. E., Nilsson, S. R., Alsio, J., Kent, B. A., Kim, & C. H., et al. (2013). The touchscreen operant platform for assessing executive function in rats and mice. *Nat. Protoc.*, 8, 1985-2005.
- Marquardt, K., Saha, M., Mishina, M., Young, J. W., Brigman, J. L. (2014). Loss of GluN2A-containing NMDA receptors impairs extra-dimensional set-shifting. *Genes Brain Behav.*, 13, 611-617.
- Marquardt, K., Sigdel, R., & Brigman, J. L. (2017). Touch-screen visual reversal learning is mediated by value encoding and signal propagation in the orbitofrontal cortex. *Neurobiology of Learning and Memory*, 139, 179-188.
- Martins, G. J., Shahrohk, M., & Powell, E.M. (2011). Genetic disruption of Met signalling impairs GABAergic striatal development and cognition. *Neuroscience*,

- 176,199-209.
- McAlonan, K., & Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behav. Neurosci.*, 146, 97-103.
- McCool, M. F., Patel, S., Talati, R., Ragozzino, M. E. (2008). Differential involvement of M1-type and M4-type muscarinic cholinergic receptors in the dorsomedial striatum in task switching. *Neurobiol. Learn. Mem.*, 89, 114-124.
- McGaughy, J., Ross, R. S., & Eichenbaum, H. (2008). Noradrenergic, but not cholinergic, deafferentation of prefrontal cortex impairs attentional set-shifting. *Neuroscience*, 153, 63-71.
- Mendenhall, W., Beaver, R. J., & Beaver, B. M. (2005). *Introduction to Probability and Statistics*. Thomson Brooks/Cole.
- Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*, 9, 100-110.
- Murray, E. A., & Gaffan, D. (2006). Prospective memory in the formation of learning sets by rhesus monkeys (*Macaca mulatta*). *J. Exp. Psychol. Anim. Behav. Process*, 32, 87-90.
- Naegeli, K. J., O'Connor, J. A., Banerjee, P. & Morilak, D. A. (2013). Effects of milnacipran on cognitive flexibility following chronic stress in rats. *Eur. J. Pharmacol.*, 703, 62-66.
- Nagahama, Y., Okina, T., Suzuki, N., Matsuzaki, S., Yamauchi, H., Nabatame, H., & Matsuda, M. (2003). Factor structure of a modified version of the Wisconsin card sorting test: an analysis of executive deficit in Alzheimer's disease and mild cognitive impairment. *Dement Geriatr Cogn Disord*, 16(2), 103-112.
- Neill, J. C., Barnes, S., Cook, S., Grayson, B., et al. (2010). Animal models of cognitive dysfunction and negative symptoms of schizophrenia: Focus on NMDA receptor antagonism. *Pharmacol. Ther.*, 128, 419-432.
- Newman, L. A., Darling, J., & McGaughy, J. (2008). Atomoxetine reverses attentional deficits produced by noradrenergic deafferentation of medial prefrontal cortex. *Psychopharmacology*, 200, 39-50.
- Ng, C. W., Noblejas, M. I., Rodefer, J. S., Smith, C. B., & Poremba, A. (2007). Double dissociation of attentional resources: prefrontal versus cingulate cortices. *J Neurosci*, 27(45), 12123-12131.
- Nieves-Martinez, E., Haynes, K., Childers, S. R., Sonntag, W. E., & Nicolle, M. M. (2012). Muscarinic receptor/G-protein coupling is reduced in the dorsomedial striatum of cognitively impaired aged rats. *Behav. Brain Res.*, 227, 258-264.
- Nikiforuk, A., Golembiowska, K., & Popik, P. (2010). Mazindol attenuates ketamine-induced cognitive deficit in the attentional set shifting task in rats. *Eur Neuropsychopharmacol*, 20, 37-48.
- Nikiforuk, A., Kos, T., Holuj, M., Potasiewicz, A., & Popik, P. (2016). Positive allosteric modulators of alpha 7 nicotinic acetylcholine receptors reverse ketamine-induced schizophrenia-like deficits in rats. *Neuropharmacology*, 101, 389-400.
- Nikiforuk, A., & Popik, P. (2011). Long-lasting cognitive deficit induced by stress is alleviated by acute administration of antidepressants. *Psychoneuroendocrinology*, 36, 28-39.
- O'Neill, M., & Brown, V. J. (2007). The effect of striatal dopamine depletion and the adenosine A2A antagonist KW-6002 on reversal learning in rats. *Neurobiol. Learn. Mem.*, 88, 75-81.
- Owen, A. M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., Lange, K. W., & Robbins, T. W. (1992). Frontostriatal cognitive deficits at different stages of Parkinson's disease. *Brain*, 115, 1727-1751.

- Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., & Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, 29(10), 993-1006.
- Perry, R. J., Watson, P., & Hodges, J. R. (2000). The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: relationship to episodic and semantic memory impairment. *Neuropsychologia*, 38(3), 252-271.
- Prentice, K. J., Gold, J. M., & Buchanan, R. W. (2008). The Wisconsin Card Sorting impairment in schizophrenia is evident in the first four trials. *Schizophr Res.*, 106, 81-87.
- Radke, A. K., Nakazawa, K., & Holmes, A. (2015). Cortical GluN2B deletion attenuates punished suppression of food reward-seeking. *Psychopharmacology*, 232, 3753-3761.
- Ragozzino, M. E. (2002). The effects of dopamine D(1) receptor blockade in the prelimbic–infralimbic areas on behavioral flexibility. *Learn. Mem.*, 9, 18-28.
- Ragozzino, M. E., Ragozzino, K. E., Mizumori, S. J., & Kesner, R. P. (2002). Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behav. Neurosci.*, 116, 105-115.
- Reid, L. S. (1953). The development of noncontinuity behavior through continuity learning. *J. Exp. Psychol.*, 46, 107-112.
- Robbins, T. W. (2007). Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. *Phil. Trans. R. Soc. B*, 362, 917-932.
- Roberts, A. C., De Salvia, M. A., Wilkinson, L. S., Collins, P., Muir, J. L., Everitt, B. J., et al. (1994). 6-Hydroxydopamine lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin Card Sort Test: possible interactions with subcortical dopamine. *J. Neurosci.*, 14, 2531-2544.
- Roberts, A. C., Robbins, T. W., & Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *Q J Exp. Psychol. B*, 40(4), 321-341.
- Rodefer, J. S., Murphy, E. R., & Baxter, M. G. (2005). PDE10A inhibition reverses subchronic PCP-induced deficits in attentional set-shifting in rats. *Eur. J. Neurosci.*, 21, 1070-1076.
- Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J. Neurol. Neurosurg. Psychiatry*, 57, 1518-1524.
- Rudebeck, P. H., & Murray, E. A. (2011). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *J. Neurosci.*, 31, 10569–10578.
- Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S., & Murray, E. A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat. Neurosci.*, 16, 1140-1145.
- Rygula, R., Clarke, H. F., Cardinal, R. N., Cockcroft, G. J., Xia J., Dalley J. W., et al. . (2015). Role of central serotonin in anticipation of rewarding and punishing outcomes: effects of selective amygdala or orbitofrontal 5-HT depletion. *Cereb. Cortex*, 25, 3064-3076.
- Scheggia, D., & Papaleo, F. (2016). An operant intra-/extra-dimensional set-shift task for mice. *Journal of Visualized Experiments*, 107, e53503.
- Schoenbaum, G., Nugent, S., Saddoris, M. P., & Gallagher, M. (2002). Teaching old rats new tricks: age related impairments in olfactory reversal learning. *Neurobiology of Aging*, 23, 555-564.
- Scott, W. A. (1962). Cognitive complexity and cognitive flexibility. *American*

- Sociological Association*, 25, 405-414.
- Seu, E., & Jentsch, J. D. (2009). Effect of acute and repeated treatment with desipramine or methylphenidate on serial reversal learning in rats. *Neuropharmacology*, 57, 665-672.
- Shiraishi, E., Suzuki, K., Harada, A., Suzuki, N., & Kimura, H. (2016). The phosphodiesterase 10A selective inhibitor TAK-063 improves cognitive functions associated with schizophrenia in rodent models. *The Journal of Pharmacology and Experimental Therapeutics*, 356, 587-595.
- Stefani, M. R., & Moghaddam B. (2005). Systemic and prefrontal cortical NMDA receptor blockade differentially affect discrimination learning and set-shift ability in rats. *Behav Neurosci*, 119, 420-428.
- Stefani, M. R., & Moghaddam B. (2006). Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *J. Neurosci.*, 26, 8810-8818.
- Stephens, D. W. (2008). Decision ecology: foraging and the ecology of animal decision making. *Cogn. Affect. Behav Neurosci.*, 8, 475-484.
- Tait, D. S., & Brown, V. J. (2008). Lesions of the basal forebrain impair reversal learning but not shifting of attentional set in rats. *Behav. Brain Res.*, 187, 100-108.
- Tait, D. S., Brown, V. J., Farovik, A., Theobald, D. E., Dalley, J. W., & Robbins, T. W. (2007). Lesions of the dorsal noradrenergic bundle impair attentional set-shifting in the rat. *Eur. J. Neurosci.*, 25, 3719-3724.
- Tait, D. S., Chase, E. A., & Brown, V. J. (2013). Tacrine improves reversal learning in older rats. *Neuropharmacology*, 73, 284-289.
- Tait, D. S., Chase, E. A., & Brown, V. J. (2014). Attentional set-shifting in rodents: a review of behavioural methods and pharmacological results. *Curr Pharm Des*, 20(31), 5046-5059.
- Tait, D. S., Marston, H. M., Shahid, M., & Brown, V. J. (2009). Asenapine restores cognitive flexibility in rats with medial prefrontal cortex lesions. *Psychopharmacology*, 202(1-3), 295-306.
- Tait, D. S., Phillips, J. M., Blackwell, A. D., & Brown, V. J. (2017). Effects of lesions of the subthalamic nucleus/zona incerta area and dorsomedial striatum on attentional set-shifting in the rat. *Neuroscience*, 345, 287-296.
- Thompson, S. M., Josey, M., Holmes, A., & Brigman, J. L. (2015). Conditional loss of GluN2B in cortex and hippocampus impairs attentional set formation. *Behav Neurosci.*, 129, 105-112.
- Torrey, E. F., Barci, B. M., Webster, M.J., Bartko, J. J., Meador-Woodruff, J. H., & Knable, M. B. (2005). Neurochemical markers for schizophrenia, bipolar disorder, and major depression in postmortem brains. *Biol Psychiatry*. 57, 252-260.
- Treasure, J., & Schmidt, U. (2013). The cognitive-interpersonal maintenance model of anorexia nervosa revisited: a summary of the evidence for cognitive, socio-emotional and interpersonal predisposing and perpetuating factors. *J. Eat. Disord.*, 1, 13-23.
- Van Eylen, L., Boets, B., Steyaert, J., Wagemans, J., & Noens, I. (2015). Executive functioning in autism spectrum disorders: influence of task and sample characteristics and relation to symptom severity. *Eur. Child Adolesc. Psychiatry* 24, 1399-1417.
- Vazey, E. M., & Aston-Jones, G. (2012). The emerging role of norepinephrine in cognitive dysfunctions of Parkinson's disease. *Front Behav Neurosci.*, 6:48.
- Waltz, J. A., & Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophr. Res.*, 93, 296-303.
- Westwood, H., Stahl, D., Mandy, W., & Tchanturia, K. (2016). The set-shifting profiles of anorexia nervosa and autism spectrum disorder using the Wisconsin Card Sorting Test: a

- systematic review and meta-analysis. *Psychological Medicine*, 46, 1809-1827.
- Wright, N. F., Vann, S. D., Aggleton, J. P., & Nelson, A. J. D. (2015). A critical role for the anterior thalamus in directing attention to task-relevant stimuli. *The Journal of Neuroscience*, 35(14), 5480-5488.
- Xia, S. (2014). The Contribution of the subthalamic nucleus to cognitive and motor *inhibition*. PhD thesis, University of St Andrews.
- Yawata, S., Yamaguchi, T., Danjo, T., Hikida, T., & Nakanishi, S. (2012). Pathway-specific control of reward learning and its flexibility via selective dopamine receptors in the nucleus accumbens. *Proc. Natl. Acad. Sci. U.S.A.*, 109, 12764-12769.
- Young, J. W., Powell, S., Risbrough, V., Marston, H., & Geyer, M. A. (2009). Using the MATRICS to guide development of a preclinical cognitive test battery for research in schizophrenia. *Pharmacol Ther.*, 122(2), 150-202.



## Appendices

### Appendix A: Robustness of rat model and human model

All the findings reported in this thesis were obtained with a specific set of pre-fixed parameter values in the Bayesian model, either for the rat or human tasks. While these findings have already shown the effectiveness and applications of the rat and human models, it would be better if the models were robust to variation in the parameters chosen for the models. Re-analysis of both rat and human data with different sets of model parameters did not reveal qualitative differences in the outcome, suggesting that the models are indeed robust. In the following, an example of the re-analysis of rat data and human data is given, respectively.

#### Lesioned rats tried more spatial trials in the ED stage than normal rats

In the rat behavioural Bayesian model, the key parameters are those in the likelihood function (Equations 2.3, 2.4, 2.6, 2.7). More specifically, the likelihood of spatial hypotheses was computed as

$$P(E_2, E_1 = 1^{st} \mid \{H = h_i, E^*\}) = \begin{cases} 0.8, & \text{if } E_2 \text{ consistent with } h_i \\ 0.1, & \text{otherwise} \end{cases}$$
$$P(E_2, E_1 = 2^{nd} \mid \{H = h_i, E^*\}) = \begin{cases} 0.05, & \text{if } E_2 \text{ consistent with } h_i \\ 0.05, & \text{otherwise} \end{cases}$$

and the likelihood of perceptual hypotheses was computed as

$$P(E_2, E_1 = 1^{st} \mid \{H = h_i, E^*\}) = \begin{cases} 0.45, & \text{if } E_2 \text{ consistent with } h_i \\ 0.05, & \text{otherwise} \end{cases}$$
$$P(E_2, E_1 = 2^{nd} \mid \{H = h_i, E^*\}) = \begin{cases} 0.45, & \text{if } E_2 \text{ consistent with } h_i \\ 0.05, & \text{otherwise} \end{cases}$$

With the above pre-fixed likelihood parameters, Section 3.5 showed that lesioned rats tried more spatial trials in the ED stage than normal rats and the difference in the number of ED trials between the two rat groups was probably mainly due to the difference in the number of spatial trials used by the two groups. To check whether such findings can be obtained by varying the Bayesian model's parameter values, I exchanged the above likelihood function with another set of parameters values, with the likelihood of spatial hypotheses being

$$P(E_2, E_1 = 1^{st} \mid \{H = h_i, E^*\}) = \begin{cases} 0.7, & \text{if } E_2 \text{ consistent with } h_i \\ 0.1, & \text{otherwise} \end{cases}$$

$$P(E_2, E_1 = 2^{nd} \mid \{H = h_i, E^*\}) = \begin{cases} 0.1, & \text{if } E_2 \text{ consistent with } h_i \\ 0.1, & \text{otherwise} \end{cases}$$

and the likelihood of perceptual hypotheses being

$$P(E_2, E_1 = 1^{st} \mid \{H = h_i, E^*\}) = \begin{cases} 0.4, & \text{if } E_2 \text{ consistent with } h_i \\ 0.1, & \text{otherwise} \end{cases}$$

$$P(E_2, E_1 = 2^{nd} \mid \{H = h_i, E^*\}) = \begin{cases} 0.4, & \text{if } E_2 \text{ consistent with } h_i \\ 0.1, & \text{otherwise} \end{cases}$$

With the new parameter values of the rat behavioural Bayesian model, I re-analysed the 47 normal and 46 lesioned rat data and obtained the same findings. In more detail, Figure A.1 (middle pair of bars) shows that lesioned rats used spatial response patterns on significantly more trials;  $m=1.57$ ,  $sd=2.44$  for the control group,  $m=6.35$ ,  $sd=5.62$  for lesion group,  $t(91)=-5.33$ ,  $p<0.001$ ,  $d=0.56$ . When I excluded these spatial trials (where the rats chose bowls using spatial response patterns) from the total trials of the ED stage for each rat, there was no significant difference between two rat groups in the remaining number of trials;  $m=11.38$ ,  $sd=4.51$  for control group,  $m=13.52$ ,  $sd=6.25$  for lesion group,  $t(91)=-1.90$ ,  $p=0.061$ , ns. Just as in Section 3.5, this suggests that the difference in ED trials between the two groups is mainly due to the difference in the number of spatial trials used by the two groups.

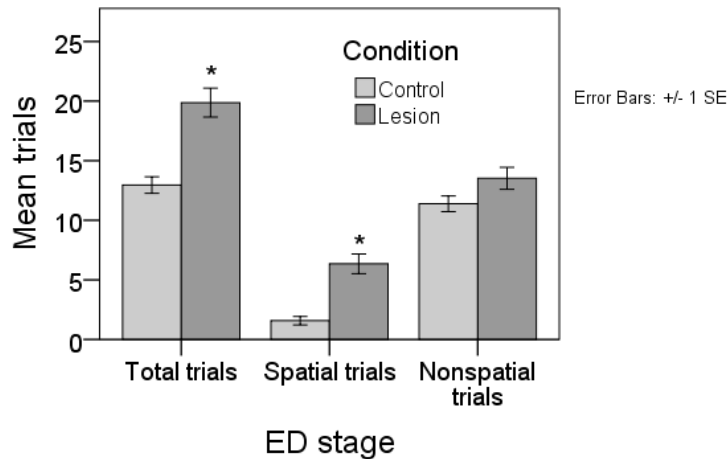


Figure A.1: number of spatial trials in ED is different between two groups.

Bayesian analysis can estimate set formation for each participant

In the human latent probabilistic model, the key parameters include the parameter  $\alpha=0.98$  in the feedback effect function (Equations 6.4-6.7) and the likelihood

parameter value 0.85 in the human behavioral reward Bayesian model (Equation 6.8). With this pre-fixed set of parameter values, Section 7.3 showed that the human behavioural reward Bayesian model could accurately estimate whether the participants had (or had not) formed an attentional set when they really did (or did not). Here I re-analysed the data by (1) only changing the feedback effect parameter from  $\alpha=0.98$  to  $\alpha=0.90$ , and (2) changing the feedback effect parameter from  $\alpha=0.98$  to  $\alpha=0.90$  and changing the likelihood parameter in the human behavioral reward Bayesian model from 0.85 to 0.90. For both sets of new parameter values, the same result was obtained, i.e., the human behavioural reward Bayesian model accurately estimated that participants had (or had not) formed an attentional set when they really had (or had not). Furthermore, when the pre-fixed threshold (0.6) used to estimate the existence of reward-irrelevant hypotheses varied from 0.55 and 0.65, again, the same result as above was obtained, supporting the view that the Bayesian analysis is robust to this threshold as well.